Investigating the Selectivity and Mechanism of Allosteric Regulation in α-IPMS Enzymes

Andrew Davies

Department of Chemistry

A thesis submitted for the degree of

Master of Science in Biochemistry

June 2015

Abstract

Enzymes are nature's wizards: balanced delicately on the margin of order and entropy, they perform chemical reactions and syntheses at rates and yields human chemists can only dream of. Many possess exquisite control mechanisms to keep the flow of metabolites through our cells precisely regulated. This work explores the regulation mechanism of α-isopropylmalate synthase (α-IPMS).

The branched-chain amino acid biosynthetic pathways in bacteria are of interest as novel antibiotic targets. α-IPMS catalyses the first committed step in the pathway to form leucine, an essential amino acid. It performs the Claisen condensation of α-ketoisovalerate (α-KIV) and acetyl coenzyme A (AcCoA) to form α-isopropylmalate (α-IPM). Almost all previously characterised α-IPMS enzymes are feedback regulated by leucine, the end-product of this pathway.

This study uses the α-IPMS enzymes from two pathogenic species, *Mycobacterium tuberculosis* and *Neisseria meningitidis* (*Mtu*IPMS and *Nme*IPMS, respectively). These enzymes are homodimeric in solution, and have a catalytic dimer of $(\beta/\alpha)_8$ barrels. This is connected via two more subdomains to a dimerised C-terminal regulatory domain, where leucine binds. The crystal structures of *Mtu*IPMS with and without leucine bound are almost identical. Thus, we do not yet fully understand the mechanisms by which leucine is

recognised, nor how the allosteric signal is conducted ~ 50 Å from the regulatory domain to the active site, and how this disrupts catalysis.

Chapter 2 explores the residues responsible for recognising and binding leucine. We use insights from the partial crystal structure of a similar enzyme in *Leptospira interrogans*, citramalate synthase (CMS). CMS catalyses a similar reaction to α-IPMS: the condensation of AcCoA and α-ketobutyrate (α-KB) to form citramalate, as the first step in isoleucine production in this organism. CMS is feedback regulated by isoleucine just as α-IPMS is regulated by leucine. CMS also shares a very similar overall structure to α-IPMS, and four conserved residues in each enzyme were identified as being responsible for binding the allosteric effector. In previous work, Tyler Clarke[1] mutated each of the four *Mtu*IPMS residues to the corresponding residue from *Li*CMS in an attempt to make an isoleucine-regulated *Mtu*IPMS. While one mutant did show an increased sensitivity to the related amino acid norvaline, none of these mutations by themselves were sufficient to create an isoleucine-sensitive *Mtu*IPMS. This work found that by using certain combinations of these mutations, we were able to create isoleucine-inhibited α-IPMS enzymes.

Dr. Wanting Jiao has been using molecular dynamics simulations to identify the residues important for allosteric signal propagation and disrupting catalysis in *Nme*IPMS . Chapter 3 details several of these residues which we have mutated, and presents the preliminary results of activity and inhibition studies on the mutant enzymes.

Chapter 4 summarises our findings and outlines the work required to further our understanding of the allosteric control systems studied here. Adapting the power of enzymes to contribute to the development of green chemistry, biosensors, and new antibiotics may prove to be one of the greatest opportunities ahead of modern chemistry.

Acknowledgements

This work is the latest step in a long journey, upon which I would have foundered many times without the kindness and brilliance of the people who have helped me along the way. I cannot possibly thank you all here, but I am grateful nonetheless.

Thank you to the teachers who inspired my love of chemistry: Neil McKeegan, Colin Moir, and Anne Drake. My professors at UC, especially Juliete Gerrard, Andy Pratt, and Emily Parker. You gave me new eyes with which to see the world. Thanks again to my supervisor, Emily. I don't know which I admire more: your seemingly boundless knowledge or your superhuman work ethic. In either case, you have been the kindest possible supervisor. Thank you for letting me come back to work with you a second time, you don't realise how much this opportunity has meant to me.

My labmates in the Parker group over the years: you all know who you are! Your experience, camaraderie, and jokes kept me going no matter how hard things were. Thanks especially to Fiona for your friendship, may your Tumblr feed never run dry! Special thanks to Nicky and Annette for your wisdom, kindness, and expertise. I would have been lost without you.

To the technical staff who keep everything running smoothly: you are the unsung heroes of our work. Thanks to Alexander Goroncy for being so good

about my never-ending mass spectrometry samples. To the friendly faces around the chemistry and biology departments whose yarns and friendship I greatly miss: Robert Stainthorpe, Wayne Mackay, Jan Wikaira, Sarah Masters, Grant Pearce, Ren Dobson, and Davey Lim, thank you!

I must also gratefully acknowledge the assistance of Freddie Mercury, Cat Stevens, Don McLean, Bob Dylan, and others who supported me greatly while writing this thesis. More seriously, thanks to Boris for emotional support, Eric for loving me, and my family who, despite the twists and turns in my journey, have never stopped believing in me.

Contents

Glossary

α-IPM

α-isopropylmalate

α-IPMS

α-isopropylmalate synthase

α-KB

α-ketobutyrate

α-KIV

α-ketoisovalerate

AcCoA

acetyl co-enzyme A

CD

Circular Dichroism

CMS

citramalate synthase

CoA

coenzyme A

DSF

differential scanning fluorimetry

DTP

4,4'-dithiodipyridine

DTT

dithiothreitol

E. coli

Escherichia coli

EDTA

ethylenediaminetetraacetic acid

HEPES

4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid

IMAC

immobilised metal affinity chromatography

IPTG

isopropyl β-D-1-thiogalactopyranoside

LB

lysogeny broth

MES

2-(N-morpholino)ethanesulfonic acid

MS

mass spectrometry

***Mtu*IPMS**

Mycobacterium tuberculosis α-IPMS

***Nme*IPMS**

Neisseria meningitidis α-IPMS

PCR

polymerase chain reaction

PDB

Protein Data Bank

SDS-PAGE

sodium dodecyl sulfate polyacrylamide gel electrophoresis

SEC

size-exclusion chromatography

TAE

tris-acetate-EDTA

TEV tobacco etch virus

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Overview

> *"The problem is so serious that it threatens the achievements of modern medicine. A post-antibiotic era—in which common infections and minor injuries can kill—is a very real possibility for the 21st century."* [2]

This chilling prediction from the World Health Organisation is a sharp reminder that our current arsenal of antibiotics is rapidly being overtaken by antibiotic resistant strains of bacteria,[3] which are developing much faster than our ability to develop new antibiotics.[4,5] While methicillin-resistant *Staphylococcus aureus* (MRSA) gains a lot of headlines in the developed world, the spectre of ancient diseases like tuberculosis still haunt the developing world (tuberculosis still kills more people than any other infectious agent except HIV/AIDS)[6] and look to be making a return to the developed world.[7] Clearly, a lot more research is needed into the mechanisms by which infectious agents achieve their virulence, and into new ways of stopping them. Understanding

the intricacies and mechanisms of key enzymes and regulatory pathways is one approach, and is the theme we explore in this work.

Enzymes responsible for catalysing the synthesis of the branched-chain amino acids (BCAAs), leucine, isoleucine, and valine, have been identified as potential drug targets,[8] and pursuant to that, the crystal structure of α-isopropyl malate synthase (α-IPMS) from *Mycobacterium tuberculosis* was obtained.[9] α-IPMS catalyses the first committed step in the leucine biosynthetic pathway. This pathway is absent in humans because we obtain all of our leucine from dietary sources, but genetic knock-out studies[10,11] show that this pathway is required for bacterial growth. The α-IPMS enzyme has proven to be rather interesting: the catalytic barrel, a fairly typical TIM or $(\beta/\alpha)_8$ barrel dimer, is allosterically regulated by a highly unusual regulatory domain so far only found in α-IPMS and a few related enzymes. While allostery has been studied for decades, it is only recently that we are beginning to untangle some of the dynamic mechanisms by which it works. By studying the unusual mechanism of allostery used in this enzyme, we may open up exciting new opportunities for drug discovery and development,[12] biosensors,[13] and provide unique insights into one of the key aspects of enzyme function.

1.2 Leucine, Isoleucine, and Valine Production: the Branched-Chain Amino Acid Biosynthetic Pathways

The branched-chain amino acids are produced by pathways that share many of the same enzymes, as can be seen in Figure 1.1. All of these pathways start with pyruvate, derived from glycolysis. Acetohydroxyacid synthase (AHAS) condenses pyruvate with either another pyruvate molecule or with α-KIV, de-

rived from threonine via threonine deaminase (TD). The condensation product is reduced and rearranged by acetohydroxyacid isomeroreductase (AHAIR). This is then dehydrated by dihydroxyacid dehydratase (DHAD) and transaminated by branched-chain aminotransferase (BCAT) to form either isoleucine or valine. Before this final step, α-ketoisovalerate (α-KIV) can be condensed with acetyl coenzyme A (AcCoA) by α-isopropylmalate synthase (α-IPMS) to form α-isopropylmalate (α-IPM). The enzyme α-isopropylmalate isomerase (α-IPMI) rearranges this to β-isopropylmalate, which is oxidised and decarboxylated by α-isopropylmalate decarboxylase (α-IPMD) to form α-ketoisocaproate. This is then transaminated by BCAT to form leucine.

1.2.1 Citramalate Pathway

In some organisms, e.g. *Leptospira interrogans*, an alternative pathway is used to form α-KB without using threonine as a starting material. Pyruvate is condensed with AcCoA by citramalate synthase (CMS) to form (R)-citramalate. The enzyme 2-methylmalate hydrolase (MMH) isomerases (R)-citramalate to form 3-methyl-D-malate, which is then oxidised by 3-methylmalate dehydrogenase (MMD) to form α-KB. This can then proceed via the isoleucine pathway. This pathway is summarised in Figure 1.2.

1.2.2 Regulation in BCAA Synthesis

The branch-points in these metabolic pathways are, as one would expect, regulated by the end-products. While variations exist between species, the general trends of regulation are summarised here. TD, the first step committing threonine to be converted to isoleucine, is inhibited by isoleucine.[14] In organisms using the citramalate pathway to product isoleucine, then CMS is inhibited by isoleucine.[15] AHAS, as the main gateway to BCAA production, is regulated

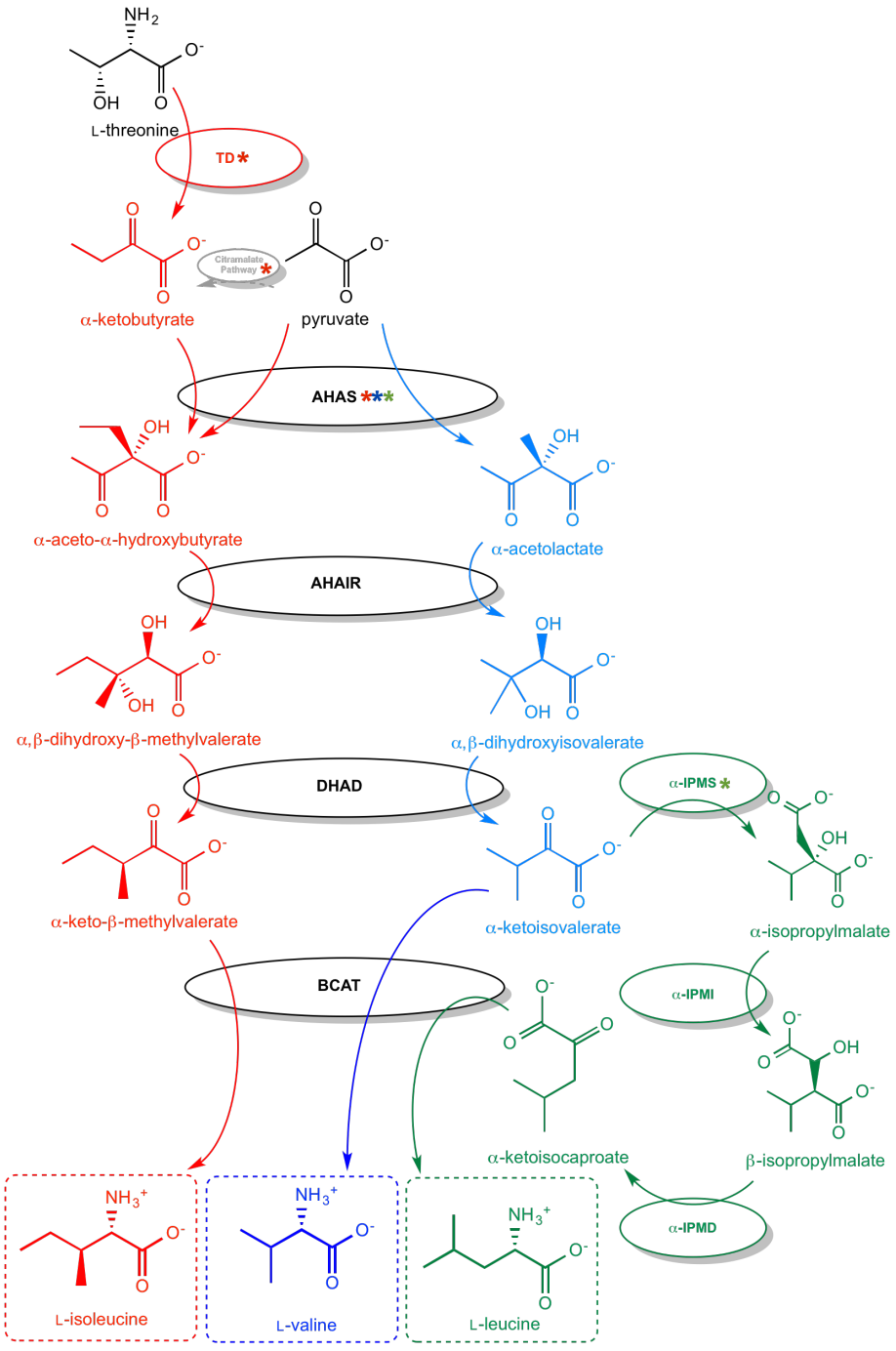


Figure 1.1: The interconnected branched-chain amino acid biosynthetic pathways. The isoleucine pathway is shown in red, valine in blue, and leucine in green. As can be seen, most of the enzymes here are used for all three pathways. The position of the citramalate pathway (delineated in Figure 1.2) is shown in grey. Asterisks denote general trends in feedback regulation by pathway end-products: red by isoleucine, blue by valine, and green by leucine.

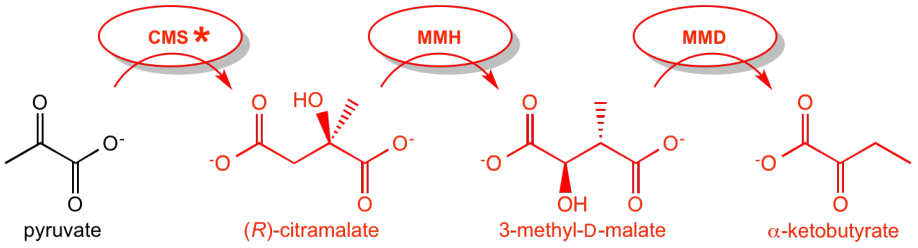


Figure 1.2: The threonine-independant citramalate pathway to form α-KB, a precursor required for isoleucine synthesis. The red asterisk indicates CMS is feedback regulated by isoleucine.

by all three: leucine, isoleucine, and valine.[16] Finally, as the branch-point for leucine synthesis, α-IPMS is inhibited by leucine.

By finely tuning the regulation at each of these metabolic control points, organisms are able to precisely regulate the flux of metabolites, ensuring their needs are met, while minimising metabolic waste. The coloured asterisks in Figure 1.1 summarise which enzymes are regulated by which metabolites.

1.3 α-Isopropylmalate Synthase

α-IPMS (EC 2.3.3.13) is at the branch point between the valine and leucine biosynthetic pathways. α-KIV either undergoes transamination by BCAT to become valine, or condensation with AcCoA by α-IPMS to go on to form leucine. α-IPMS is therefore feedback regulated by leucine to ensure sufficient levels of both leucine and valine are produced in the cell.

α-IPMS is part of a larger enzyme family of α-ketoacid: AcCoA *re*-aldolases (KARA). These catalyse the Claisen condensation of an α-keto acid with AcCoA, and members of this class of enzymes include α-IPMS, citramalate synthase (CMS), homocitrate synthase (HCS), *re*-citrate synthase (RCS), and methylthioalkylmalate synthase (MAMS).[17–20] These enzymes generally require a di-

valent metal ion for catalysis, which orientates and activates the *alpha*-carbonyl for nucleophilic attack.

The *leuA* gene (most common designation for the gene encoding α-IPMS) exhibits variable-number tandem repeats (VNTR) of a 57-nucleotide sequence, naturally present between two and twenty-one times[21] (*Mtu*IPMS strain H37Ra used in this study has two repeats). Deletion of this sequence has no significant effect on catalysis.[22]

A large body of work is based on *M. tuberculosis* α-IPMS (*Mtu*IPMS) due to its desirability as a drug target, and as until recently this was the only α-IPMS with a full-length crystal structure available.

1.3.1 *Mtu*IPMS Structure

Full length crystal structures have been found for *Mtu*IPMS[9] and *L. biflexa* α-IPMS (short form, *leuA2*).[23] The *L. biflexa* enzyme lacks a regulatory domain, and will be discussed later. The *Mtu*IPMS structure is composed of an asymmetric homodimer, depicted in Figure 1.3.

Catalytic Domain

The catalytic domain is a dimer of $(\beta/\alpha)_8$ barrels. Subdomain I wraps around to cap the active site of the other chain with an α-helix, contributing necessary residues for catalysis. A short linker connects subdomains I and II, the latter being a trio of α-helices, which support the dimerised $(\beta\beta\beta\alpha)_2$ regulatory domain. Leucine can be found (PDB: 3FIG) bound to each half of the dimer, near the dimer interface, as shown in Figure 1.4. When leucine is absent, the structure is practically identical (aligning C_α atoms in 1SR9 and 3FIG - leucine absent and bound, respectively - gives RMS = 0.23 Å).

Active Site & Subdomain I

The active site is located at the C-terminal end of each catalytic barrel. The residues Asp-81, His285, and His-287 bind a divalent metal ion (i.e. Zn^{2+}), responsible for orientating α-KIV by coordinating the two carbonyl groups, and activating the α-carbonyl for nucleophilic attack. α-KIV is further stabilised by hydrogen bonds to Arg-80 and Thr-254. These details are shown in Figure 1.5.

Subdomain II

Subdomain II looks like a hinge, and occupies quite a different orientation with respect to the catalytic domain in each chain (see Figure 1.6). The reason for this asymmetry is unclear. It appears to play a role in conducting the allosteric signal from the regulatory domain to the active site.[24] Truncations that disrupt subdomain II inactivate the enzyme,[23,25] and it has been argued[23] that subdomain II may play a crucial role in coordinating AcCoA for catalysis. This is supported by the findings of Zhang *et al.*[23] that partially truncating *L. biflexa* α-IPMS (short form) to disrupt subdomain II does not affect α-ketobutryate binding, but doubles the K_m for AcCoA and drastically lowers the catalytic rate.

Regulatory Domain

It has been shown that the regulatory domain is not required for catalysis,[23] but it is the site where leucine binds to effect allosteric regulation. This was depicted in Figure 1.4, and will be discussed further in the next chapter.

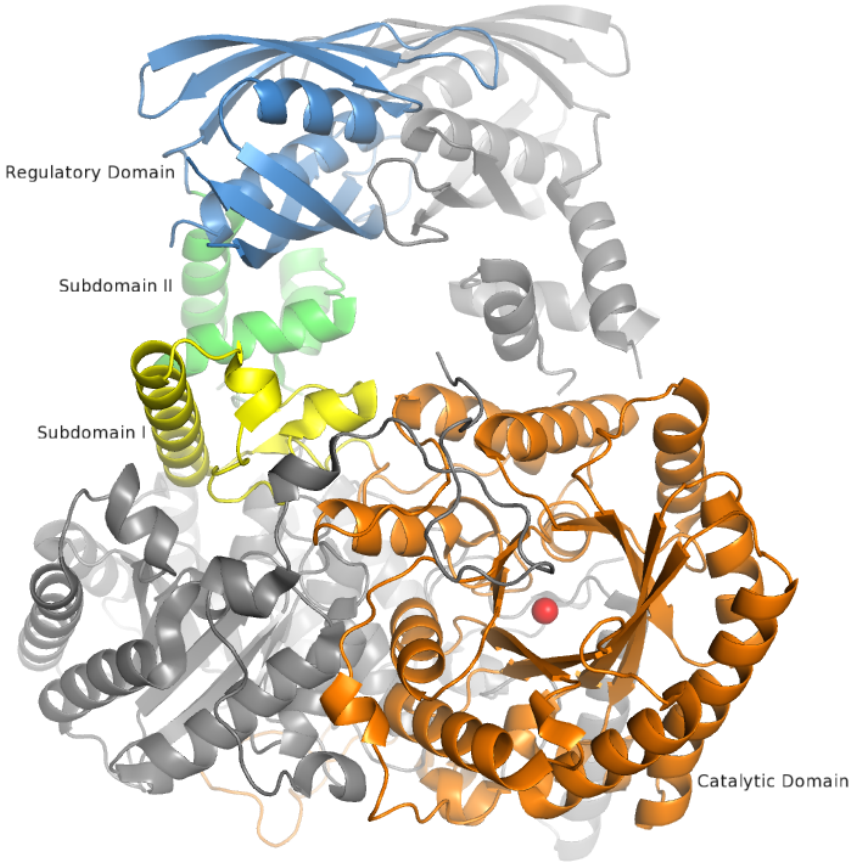


Figure 1.3: *Mtu*IPMS crystal structure (PDB ID: 3FIG). Chain A is in grey, chain B is coloured according to domain. Zn^{2+} ion shown in the active site of chain B as a red sphere. Some flexible loops are not shown due to missing electron density.

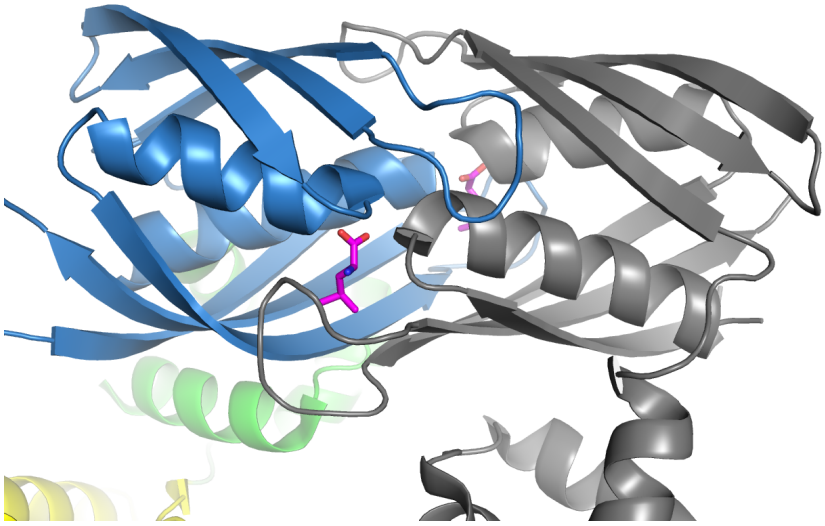


Figure 1.4: *Mtu*IPMS crystal structure (PDB ID: 3FIG) showing close-up of the regulatory domain. Bound leucine molecules are shown in pink.

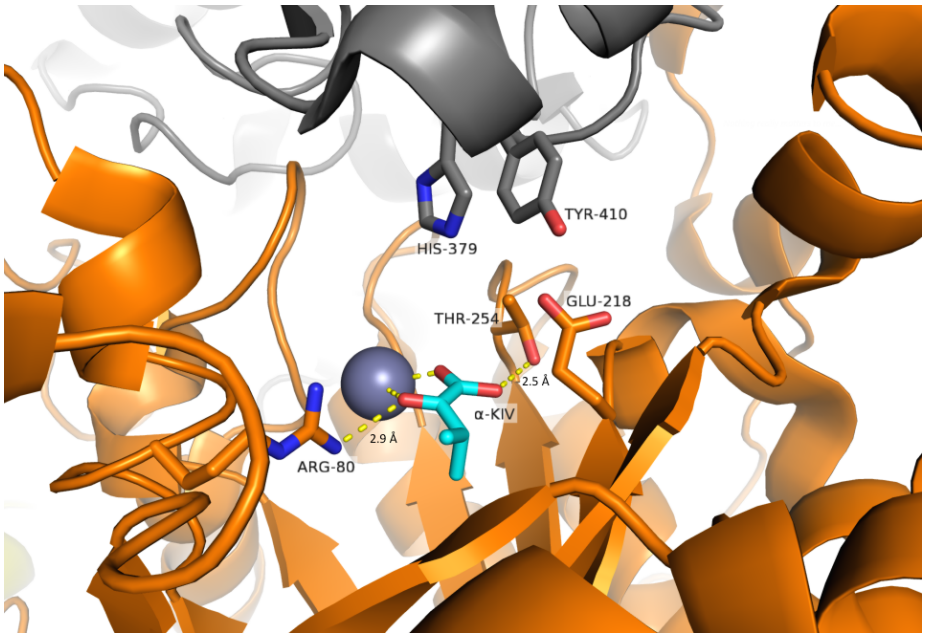


Figure 1.5: *Mtu*IPMS active site. Zn^{2+} depicted as a sphere. Note the contribution of His379 and Tyr410 from the other chain, via subunit I (grey).

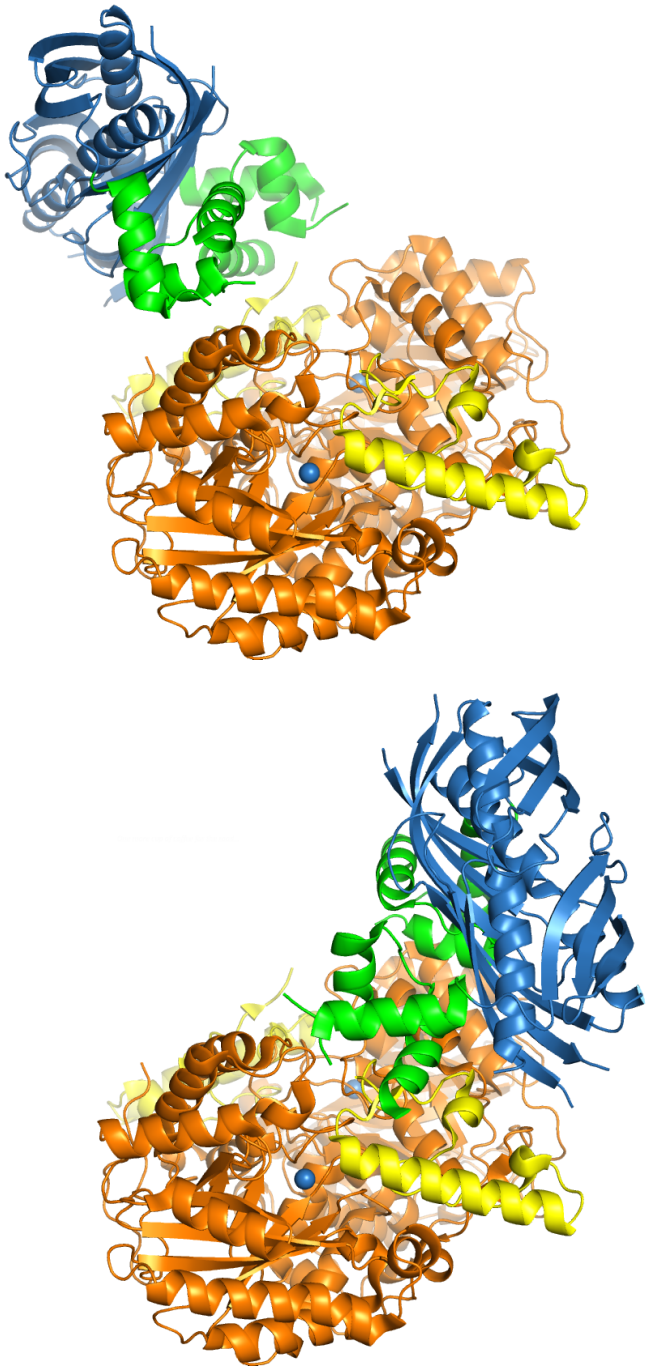


Figure 1.6: Subdomain II appears to act as a hinge between the catalytic barrel and the regulatory domain. The view from each side of the enzyme is shown here to demonstrate the asymmetrical arrangement of subdomain II and the regulatory domain with respect to the catalytic domain. Missing electron density from flexible loops is more obvious here than in Figure 1.3.

1.3.2 *Lb*IPMS

The crystal structures (PDB ID: 4OV4, 4OV9) of *Lb*IPMS align closely with that of *Mtu*IPMS, as shown in Figure 1.7. The obvious exception is subdomain II: as *Lb*IPMS lacks the constraints of the *Mtu*IPMS regulatory dimer, subdomains II adopt a different conformation. They are still asymmetrically positioned, just as in *Mtu*IPMS. The similarity between the structures reinforces the generalisations about α-IPMS we have drawn from *Mtu*IPMS.

1.3.3 Catalysis

*Mtu*IPMS is a metalloenzyme, catalysing a Claisen condensation between α-KIV and AcCoA under slightly basic conditions. While a divalent metal ion is required for catalysis, frequently monovalent cations (i.e. K^+, Na^+, and $NH_4{}^+$) are activating in several α-IPMS enzymes.[26–28] As these don't appear in any crystal structures, we are unable to determine their role. They may play a role in recruiting the divalent metal to the active site.[27]

De Carvalho *et al.*[27] reported K_m values for α-KIV and AcCoA of 12 and 136 mM respectively, with a turnover rate of 3.5 s^{-1}. A comparison with various other α-IPMS enzymes is shown in Table 1.1. *Mtu*IPMS has been shown to be able to catalyse AcCoA hydrolysis without α-KIV present. The K_m is similar (160 μM) but with negligible turnover (0.03 s^{-1}).

*Mtu*IPMS has been shown to be somewhat promiscuous with respect to α-keto acids. For instance, it is able to use pyruvate, α-KB, α-ketovalerate, α-ketoisocaproate, and α-ketocaproate (shown in Figure 1.8).[29,30] Not surprisingly, some orthologs are inhibited by α-ketoisocaproate[31] and α-ketovalerate[30] and similar structures, presumably due to competitively binding to the active site.

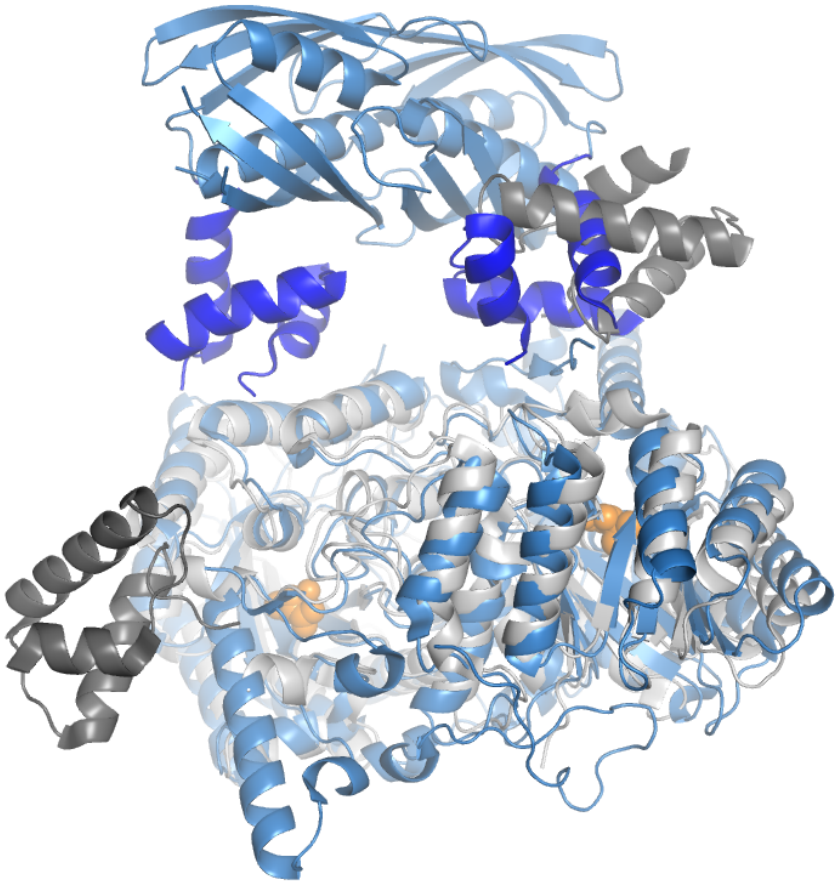


Figure 1.7: Overlay of *Lb*IPMS (4OV4) with *Mtu*IPMS (3FIG). Catalytic domains align with RMS = 1.71 Å. *Mtu*IPMS is shown in sky blue, with subdomain II in dark blue. *Lb*IPMS is shown in grey, with subdomain II in darker grey. Active site molecules depicted as orange spheres.

Mechanism

The mechanism proposed by De Carvalho *et al.*[31] is shown in Figure 1.9. A divalent metal ion coordinates the carbonyl groups of α-KIV, activating the α-carbonyl for nucleophilic attack. A basic residue abstracts a proton from the terminal methyl group of AcCoA, which attacks the carbonyl by a concerted mechanism. A second base abstracts a proton from water, providing a nucleophile to hydrolyse AcCoA, releasing CoA and α-isopropylmalate.

Table 1.1: Comparison of basic kinetic parameters for several α-IPMS enzymes. Not all studies reported a k_{cat} value.

Organism	K_m α-KIV (μM)	K_m AcCoA (μM)	k_{cat} (s^{-1})
M. tuberculosis[27]	12	136	3.5
Neisseria meningitidis[32]	30	35	13
*Arabidopsis thaliana**[30]	304	45	2.4
N. crassa[26]	10	25	–
S. cerevisiae[28]	16	9	–
Salmonella typhimurium[33]	60	200	–

* Isozyme 1

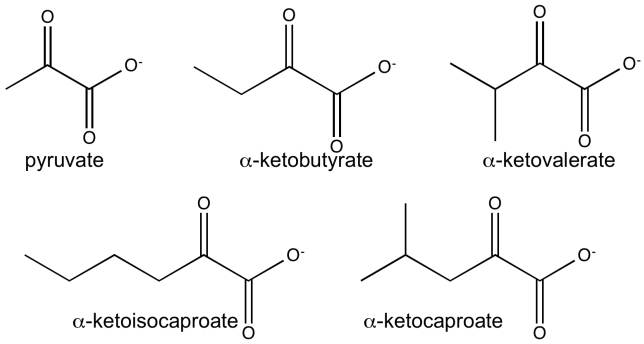


Figure 1.8: Some alternative substrates for *Mtu*IPMS.

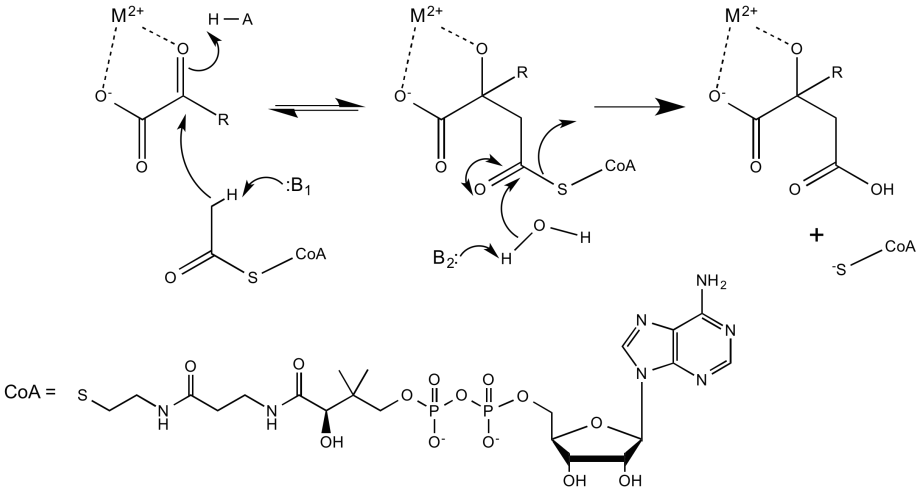


Figure 1.9: Mechanism adapted from De Carvalho *et al.*[31] B_1 and B_2 are unidentified basic residues, and H-A an acidic residue in the active site. A generalised α-keto acid is shown here rather than α-KIV.

1.3.4 Regulation

α-IPMS is allosterically feedback-regulated by L-leucine, but the mechanism by which the allosteric signal is propagated approximately 50 Å from allosteric to catalytic site is not fully understood. Our understanding of allostery has gradually evolved since the 1960's, when the theories of Monod *et al.*[34] and Koshland *et al.*[35] were first proposed. Each theory suggested that enzymes could exist in one of two states, a "relaxed" state: able to bind substrate and conduct catalysis; and a "tense" state (caused by inhibitor binding), that prevents substrate binding and catalysis. While these theories provide a reasonable explanation of the gross conformational shifts seen in some systems, it is now recognised that enzymes are necessarily dynamic, flexible systems. Rather than two rigid conformations, enzymes occupy a range or "ensemble" of conformations that shift according to physical and chemical conditions. Upon binding an allosteric

effector, the ensemble shifts to a different range of conformations, which results in activation or inhibition of catalysis.[36,37]

As leucine binding has almost no effect on the crystal structure of *Mtu*IPMS, implying the nature of allosteric regulation must be through altered enzyme dynamics, rather than gross changes in structure. Frantom *et al.*[24] probed the allosteric network in *Mtu*IPMS by amide hydrogen/deuterium Exchange (HDX) experiments. Enzyme was incubated in D_2O before being subjected to proteolysis and mass spectroscopy. Experimental conditions could then be altered, and changes in deuterium uptake would reflect any change in solvent exposure or protection certain regions of the protein experienced.

As would be expected, several peptide fragments originating near the allosteric site showed a decrease in deuteration upon leucine binding (see Figure 1.10). However, one fragment (residues 78-87) from the catalytic site also showed decreased deuteration: this region includes several conserved residues, including Arg-80, which helps stabilise α-KIV. The same fragment showed decreased deuteration in the enzyme mutant Y410F.

As can be seen in Figure 1.5, Tyr410 is present on a loop in subdomain I, and is thought to orientate His379 for catalysis.[9] Tyr410 is, in turn, orientated by hydrogen bonds from its phenolic moiety to Gln176 and Glu218. Mutating to phenylalanine disrupts these contacts, and severely compromises catalysis. The catalytic rate of this mutant is similar to the fully inhibited wild-type, suggesting this mutation mimics the inhibited wild-type enzyme. While this gives us clues as to the proximal mechanism of inhibition, it reveals little about the transmission of the allosteric signal to the active site.

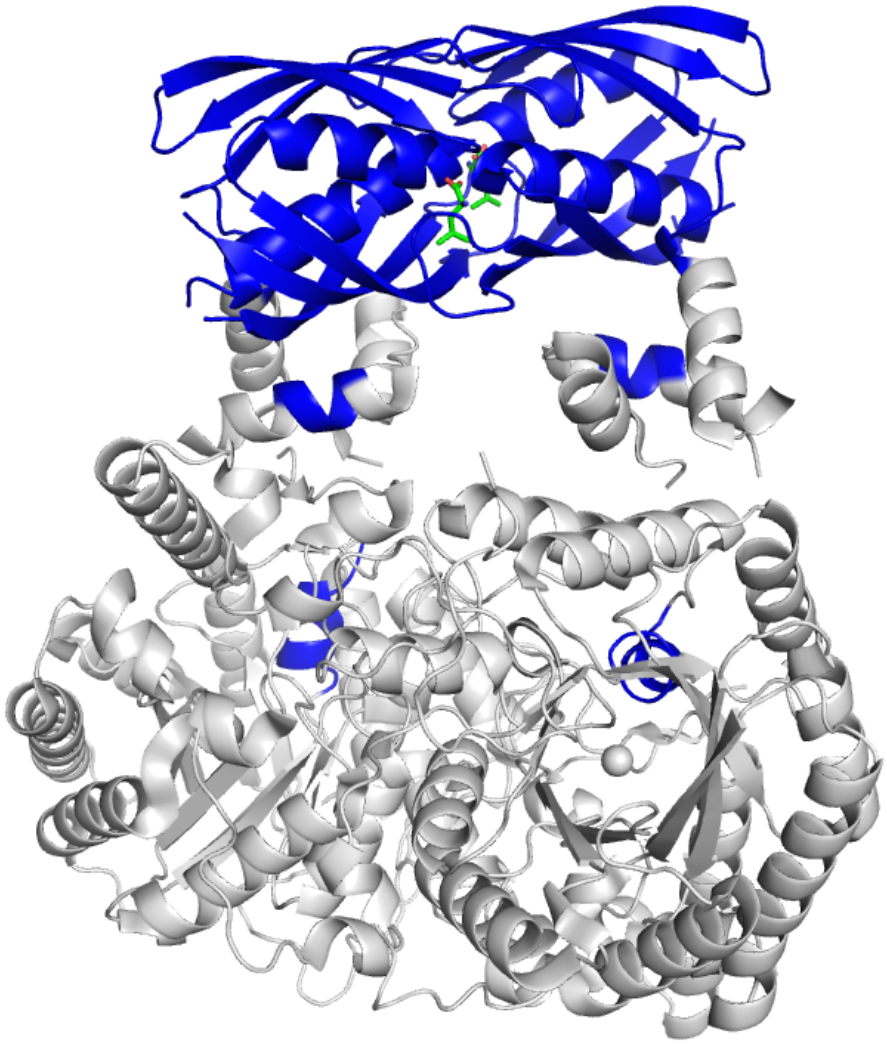


Figure 1.10: HDX experiments show areas of decreased solvent exposure in *Mtu*IPMS upon leucine binding (regions shaded blue). Leucine is shown bound in the allosteric site (green).

1.4 Citramalate Synthase

Citramalate synthase (CMS) (EC 2.3.1.182) is the first step in an alternative pathway for producing α-KB from pyruvate rather than threonine. It has only been characterised in a handful of organisms.[38–43] To date, Swiss-Prot lists only 13 other CMS sequences.[44]

CMS catalyses the first step in the citramalate pathway shown in Figure 1.2, the Claisen-condensation of two pyruvate molecules to form (*R*)-citramalate. It is feedback-regulated by isoleucine. Available kinetic parameters are summarised in Table 1.2.

Table 1.2: Kinetic parameters available for CMS enzymes.

Organism	K_m pyruvate (μM)	K_m AcCoA (μM)	k_{cat} (s^{-1})
L. interrogans	60	1100	10
Methanococcus jannaschii	180	303	0.36
Rhodospirillum rubrum	–	3200	–

Partial crystal structures exist for *L. interrogans* CMS (*Li*CMS) (PDB ID: 3F6G - regulatory domain, 3BLI - catalytic domain).[15] *L. interrogans* is the causative agent in "rat-catcher's fever", a zoonotic infection primarily contracted by contact with soil or water contaminated with animal urine. The crystal structures align well with the *leuA* structure, as seen in Figure 1.11, with RMS = 2.5 Å (regulatory domains) and 1.7 Å (catalytic domains). Isoleucine is bound in the same pocket that binds leucine in *Mtu*IPMS (show in Figure 1.12). While no structures exist of *Mtu*IPMS with AcCoA bound, it is present in *Li*CMS (PDB ID: 3BLI). This is shown in Figure 1.13.

Figure 1.11: Partial crystal structures of *Li*CMS (blue) aligned to *Mtu*IPMS (grey) with RMS $\leq$ 2.5Å. PDB IDs: 3F6G (regulatory domain) and 3BLI (catalytic domain).

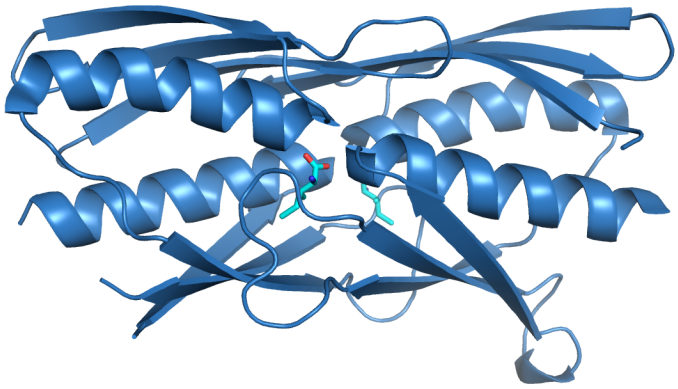


Figure 1.12: Isoleucine-binding pocket with isoleucine bound (lighter blue). PDB ID: 3F6G.

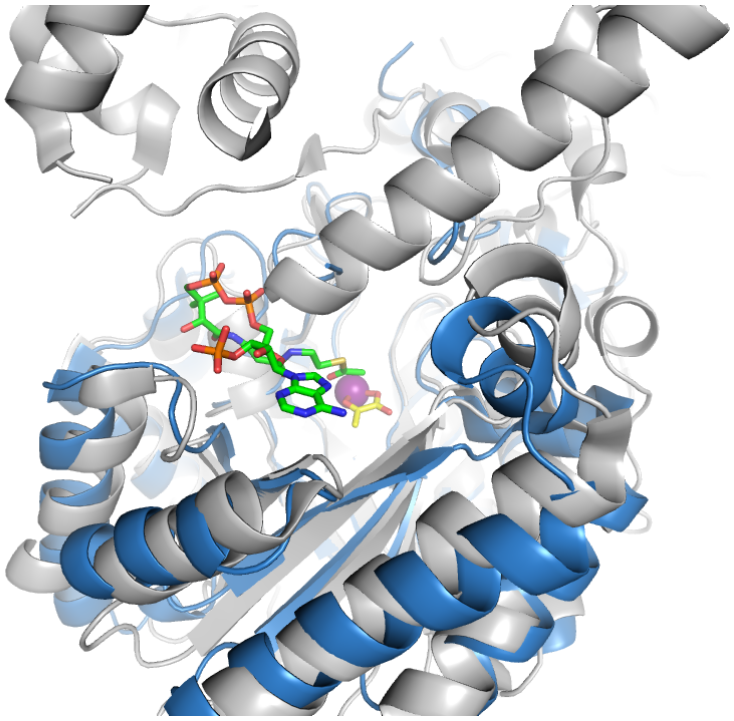


Figure 1.13: AcCoA binding in *Li*CMS (blue), pyruvate (yellow), and Zn^{2+} (purple). Overlaid is *Mtu*IPMS (grey), indicating where AcCoA may bind in that enzyme. PDB ID: 3BLI.

1.5 Thesis Objective

The objective of this thesis was to explore the selectivity and underlying mechanism of allostery in α-IPMS enzymes. One aim was to investigate the residues binding the allosteric effector in *Mtu*IPMS, and gain a better understanding of how they cooperate to mediate effector recognition. The other main aim was to investigate some of the residues highlighted by molecular dynamics simulations, and determine whether they play a role in transmitting the allosteric signal from the regulatory domain to the catalytic domain.

The enzymes used in this study, *Mtu*IPMS and *Nme*IPMS were chosen as they were already reasonably well-characterised in our lab, and in the literature. Enzymes from these organisms are of general scientific interest due to the pathogenicity of their host organisms, and by pursuing a range of different angles of research, we hope to gain insights that may be useful in curbing their virulence.

Chapter 2

Altering Allosteric Specificity of *Mycobacterium tuberculosis* α-IPMS through Rational Design

2.1 Overview

*Mtu*IPMS is allosterically inhibited by leucine, but not significantly by chemical analogues such as isoleucine. The goal of this project was to mutate the allosteric site of *Mtu*IPMS so that it would bind and be inhibited by isoleucine or similar amino acids.

Using insights from X-ray crystallography and building on previous work by Tyler Clarke,[1] multiple-point site-directed mutagenesis (SDM) was used to create several candidate enzymes for analysis.

The enzymes were expressed and purified. A circular dichroism (CD) spectrum was obtained for each enzyme and compared with the wild-type to the secondary structure. Differential scanning fluorimetry (DSF) was trialled as a first-pass screen for inhibitory activity of various amino acids. Finally, enzyme inhibition studies were carried out to test whether the allosteric specificity of *Mtu*IPMS had changed as expected.

2.2 Background

*Mtu*IPMS and *Li*CMS catalyse a very similar reaction: an aldol reaction between AcCoA and a small α-keto acid. Although they show only around 34% sequence similarity[a], crystal structure overlays show a high degree of similarity. The crystal structure of *Mtu*IPMS (PDB ID: 3FIG) aligns to the crystal structures of *Li*CMS with a RMS of 1.95 Å for catalytic domain (PDB ID: 3BLE) Cα atoms, and 3.03 Å for regulatory domain Cα atoms (PDB ID: 3F6G). The resolution of these structures were 2.30 Å, 2.00 Å, and 2.00 Å respectively. *Mtu*IPMS and *Li*CMS are both allosterically inhibited by the end-products of their metabolic pathways, leucine and isoleucine respectively.

In the characterisation of *Li*CMS by Zhang *et al.*,[15] it was found that a single mutation (V468A) was sufficient to result in allosteric inhibition with leucine. Interestingly, this mutant was twice as sensitive to isoleucine as the wild-type enzyme, but it had gained some measure of leucine sensitivity, as can be seen from the IC_{50} values displayed in Table 2.1.

[a]Pairwise alignment from www.ebi.ac.uk using a BLOSUM62 matrix

Table 2.1: Inhibition values for *Li*CMS V468A, adapted from Zhang *et al.*[15]

*Li*CMS	IC_{50} Ile	IC_{50} Leu
wild-type	86 ± 20	—
V468A	41 ± 8	343 ± 41

Attempting to produce the reciprocal result in *Mtu*IPMS, Clarke[1] identified four residues thought to be key for leucine binding in *Mtu*IPMS. These residues, Val551, Ala558, Ala567, and Ile627 were mutated to their corresponding CMS residues and the resultant enzymes were characterised. Interestingly, none of these mutations introduced any sensitivity to isoleucine, but A558I and I627A each showed a marked reduction in leucine sensitivity. A567V (the mutation corresponding to *Li*CMS V468A) was still sensitive towards leucine, but showed an increased sensitivity towards norvaline. This project tests a rational combination of these mutations to see whether *Mtu*IPMS can be made sensitive to isoleucine by a combination of two or more mutations.

2.3 Residues Important for Allostery

Table 2.2: Distances between nearest points of allosteric site residues and allosteric effector in crystal structures.

*Mtu*IPMS		*Li*CMS	
Residue	Avg. distance to leucine (Å)	Residue	Avg. distance to isoleucine (Å)
Val-551	4.3	Leu-451	4.4
Ala-558	4.1	Ile-458	4.1
Ala-567	4.0	Val-468	3.8
Ile-627	3.8	Gln-495	3.6

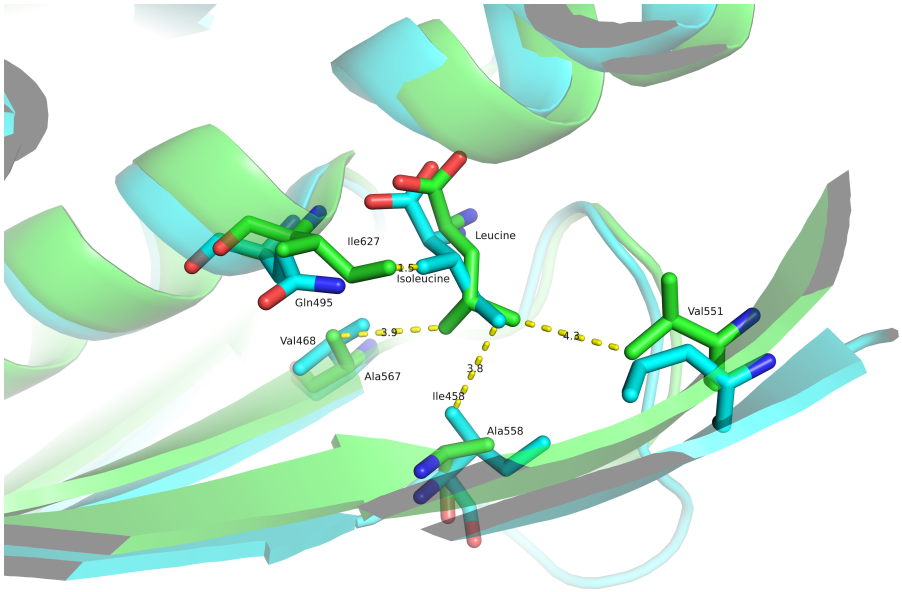


Figure 2.1: Overlay of allosteric sites of *Mtu*IPMS (green) & *Li*CMS (blue). Allosteric effector molecules and important residues depicted and labelled. PDB structures (3FIG & 3F6G, respectively) aligned by Cα positions. Some representative distances are shown.

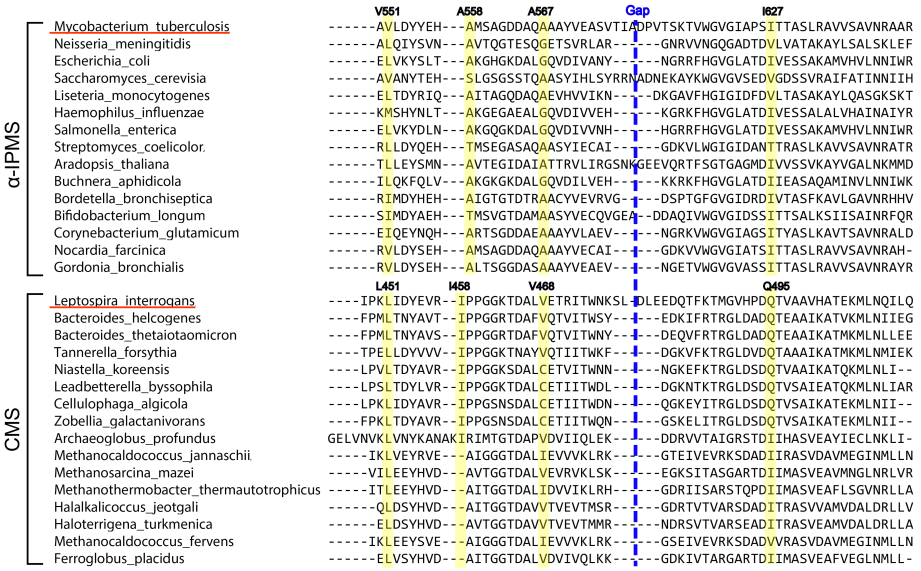


Figure 2.2: MSA of allosteric binding site residues in α-IPMS and CMS Enzymes. MSA constructed using Jalview/ClustalO [45,46]

A close examination of the crystal structure of *Mtu*IPMS overlaid with the allosteric domain of *Li*CMS reveals interesting details about the allosteric binding site of these enzymes (Figure 2.1). Table 2.2 lists the main residues thought to be responsible for allosteric effector binding from *Mtu*IPMS and *Li*CMS. The multiple sequence alignment (MSA) in Figure 2.2 shows these residues are reasonably well-conserved across various α-IPMS enzymes.

The lack of absolute conservation implies that either than absolute specificity is not important, or that the residues involved in mediating allosteric effector recognition work cooperatively rather than through individual, specific interactions. For instance, Fox *et al.* found in a directed evolution experiment that combinations of mutations lead to a positive outcome that had no effect by themselves. While none of the mutants tested by Clarke showed any isoleucine

sensitivity, this may be because a combination of mutations are required to sufficiently alter the allosteric pocket.

The kinetic assays revealed no significant changes to K_m or k_{cat} for any of the mutants, but the inhibition assays showed some enzymes had lost their sensitivity to leucine. None of the enzymes were inhibited by isoleucine, but A567V developed a strong sensitivity towards norvaline, while suffering a small decrease in sensitivity towards leucine. Presumably the mutant valine sidechain adds just enough steric bulk to discourage the isopropyl terminus of leucine's sidechain, while providing enough extra hydrophobic interactions to make binding to norvaline's propyl sidechain more favourable. As isoleucine shares this unsubstituted propyl group, it may prove to be one of the more important residues for isoleucine binding. See Figure 2.3 to compare the sidechains of leucine and isoleucine.

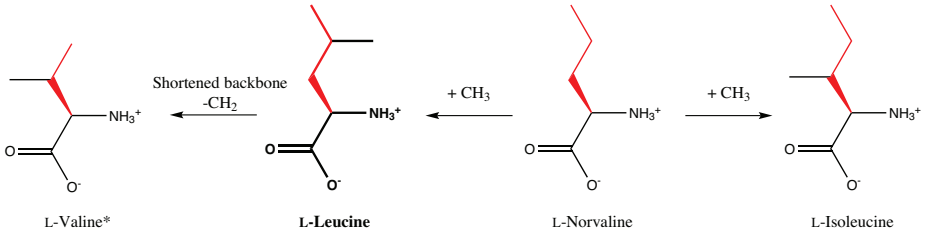


Figure 2.3: Amino acids tested as inhibitors in this study. The sidechain of norvaline is coloured red, and is shown (substituted or truncated) in the other amino acids, showing the structural similarity explored in this study. *Valine was not tested in the previous study.

Based on the crystal structure, we would expect I627A to remove the steric hindrance preventing accommodation of isoleucine's β-methyl group. The study found that this mutation also removes inhibition by leucine and norvaline. Presumably some of the hydrophobic interactions stabilising the binding of leucine or isoleucine have also been removed. The other two mutants (A558I and V551L) strongly decrease leucine sensitivity, and weaken that of norvaline.

Their positions in the crystal structure suggest a role in regulating the amount of space available for aliphatic sidechains to be accommodated, which would explain why they reduce sensitivity to leucine, but not necessarily norvaline. Their role in shaping isoleucine sensitivity will be further explored in this project.

In preparing for this work, the DSF results obtained by Clarke were carefully analysed. DSF measures the thermostabilising effect produced when small molecules bind to a protein,[47] and this was used in an attemot to predict the allosteric properties of the enzymes. Figure 2.4 shows the predictive value of DSF in the previous study was, at most, rather limited ($R^2 = 0.246$). This is not entirely surprising as we can only infer that the amino acids are binding at the allosteric site: they may in fact have some non-specific binding properties. An example of this possibility would be the use of arginine in the refolding misfolded proteins,[48] (although admittedly at much higher concentrations). We must also consider that there will be both enthalpic and entropic factors affecting the binding of ligand molecules and the ability of these to stabilise large, dynamic systems such as proteins. DSF will be further explored in this study, with a larger sample size, to explore whether it provides a good indication of allosteric properties of *Mtu*IPMS.

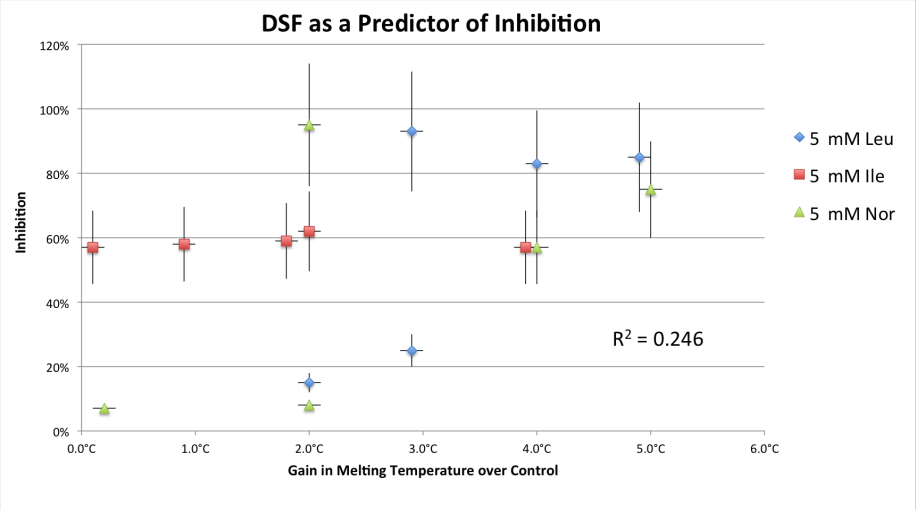


Figure 2.4: DSF was a reasonably poor predictor of which amino acids would be allosteric effectors for previously studied *Mtu*IPMS mutants. Data is shown for 5 mM concentrations, lower values had a similar effect (data not shown). DSF was carried out on 25 μL samples containing 160 μg/mL protein in BTP buffer (pH 8.0).

2.4 Generation of Mutants

As A567V had shown some promise in opening up the allosteric profile of *Mtu*IPMS it was treated as a "base" mutation, to which was added various combinations of the other mutations as delineated in in Table 2.3.

Table 2.3: Table of planned mutations

Mutation	Mutant, designated **"Mtu-..."**						
	1	2	3	4	5	6	7
A567V	X	X	X	X	X	X	X
A558I	X		X		X		X
V551L		X	X			X	X
I627A				X	X	X	X

The first step was to obtain a pure sample of pProEx-HTa-LeuA plasmid, containing the *Mtu*IPMS (*leuA*) gene along with several other features to aid purification. A T7 promotor is present upstream of the gene to allow induction of overexpression with isopropyl β-D-1-thiogalactopyranoside (IPTG). An N-terminal $(His)_6$-tag simplifies purification, and a tobacco etch virus (TEV) protease recognition site allows tag removal. A gene encoding for ampicillin-resistance acts as a selection marker. A small overnight culture of bacteria carrying the plasmid was processed using a commercially available plasmid-extraction kit to obtain a sample of the plasmid. This was checked by agarose gel electrophoresis for size and purity, as shown in Figure 2.5.

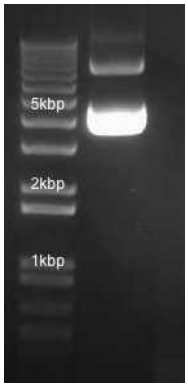


Figure 2.5: Purified plasmid, with supercoiled, linear, and nicked forms clearly seen.

As the project required multiple steps of SDM to produce the required mutants, obtaining pure template DNA was crucial for effective SDM reactions. Initial SDM attempts were repeated many times, varying a wide variety of conditions: but consistently failed. This was eventually blamed on cellular RNA co-purifying with and contaminating the extracted plasmid[b]. The RNA had initially been missed, due to longer electrophoresis running times leading to the RNA running off the gel. The presence of RNA would significantly

[b] Annette Stewart, personal communication, August 2014

inflate the nucleic acid concentration estimated by OD_{260}, leading to insufficient template being used in the SDM reactions. Changing plasmid-extraction kit suppliers resulted in much purer plasmid extracts (no RNA visible on agarose gels), which allowed further SDM to proceed smoothly. SDM products were digested with DpnI restriction enzyme to remove methylated (template) DNA, leaving the newly synthesised, unmethylated mutant DNA intact. This was used to transform chemically competent *E. coli* cells, which were grown overnight and several colonies picked for plasmid extraction. Extracted plasmids were typically subjected to PCR using pProEx sequencing primers, and both plasmids and PCR products analysed by agarose gel electrophoresis to ensure correct size (see Figure 2.6). Plasmids were submitted for DNA sequencing analysis to confirm the mutation and integrity of the gene. A sample of a plasmid with the correctly mutated sequence was used to transform chemically competent *E. coli* expression strain cells, which were stored as glycerol stocks at -80°C until required for growth and protein expression.

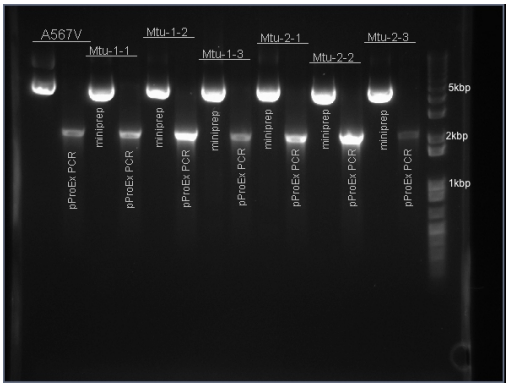
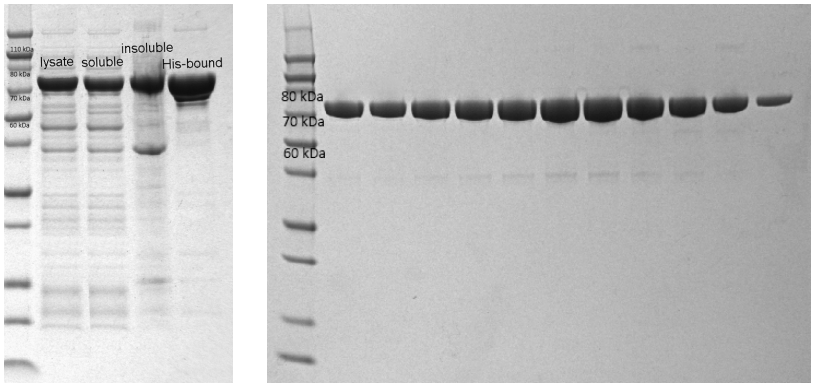


Figure 2.6: Checking mutant plasmids and their PCR products.

2.5 Protein Expression and Purification

Wild-type and mutant *Mtu*IPMS enzymes were grown and expressed as per the methods described in Section 5.4. The various stages of purification can be seen on SDS-PAGE gels in Figure 2.7. The following sections outline this process.



(a) HisTrap (b) Size-Exclusion Chromatography

Figure 2.7: Mtu-7 purification visualised by SDS-PAGE.

2.5.1 Protein Extraction

Fresh or thawed cell pellets were resuspended and lysed by sonication on ice with Benzonase® Nuclease (Sigma) to break down nucleic acids. Lysate was clarified by centrifugation and filtration, and the insoluble pellet resuspended in 8 M urea for electrophoresis. The lysate and soluble fractions in Figure 2.7a show clear overexpression of the tagged *Mtu*IPMS enzyme at 73 kDa, but the insoluble fraction still contains a certain amount of protein, suggesting incomplete lysis.

2.5.2 Immobilised Metal Affinity Chromatography (IMAC)

The clarified soluble fraction was loaded onto a HisTrap (GE Healthcare) IMAC column. These columns coordinate Ni^{2+} ions, but leave several coordination

sites free. These sites are bound by imidazole moieties on polyhistidine tags (His-tags), causing highly specific binding of tagged proteins to the column. An imidazole gradient is used to elute the protein, which is monitored by absorption at 280 nm. The protein elutes reasonably pure, as can be seen in the pooled fractions, gelled in the final lane of Figure 2.7a. A desalting column was then used to remove the imidazole, to prevent it interfering with the following step.

2.5.3 TEV Protease Treatment

To remove any interference in enzyme function from the His-tag, the tag was cleaved by TEV protease at a recognition sequence designed into the short linker between the polyhistidine section and the beginning of the *Mtu*IPMS sequence. This sequence is described in Figure 2.8.

MHHHHHHGKPIPNPLLGLDSTENLYFQGA

Figure 2.8: The hexahistidine section (red) binds to the IMAC column, followed by a linker region (orange), and a TEV protease recognition site (underlined). The GA residues (blue) after the recognition site remain attached to the protein.

Approximately 1 mg TEV protease was added per 10-100 mg tagged protein, along with DTT and EDTA. A brief (<60 mins) benchtop incubation step was followed by overnight incubation at 4°C. The sample was desalted again to remove the DTT and EDTA, before being subjected to a second IMAC step. This bound the His-tagged TEV protease enzyme and the cleaved His-tag, allowing the cleaved protein to elute.

2.5.4 Size-Exclusion Chromatography (SEC)

A final polishing step was carried out via SEC. The enzyme eluted as a single peak, and fractions were examined by SDS-PAGE (Figure 2.7b). The mass inferred from the gel ($\approx$ 70 kDa) matches the predicted protein size. All fractions

visually judged to be >95% pure were pooled and concentrated to >1 mg/mL, and flash-frozen in liquid nitrogen before being stored at -80°C.

2.6 Physical Characterisation

2.6.1 Mass Spectrometry

Mass spectrometry proved difficult for many of the samples prepared in this report, despite cleaning up samples on an analytical SEC column and repeated desalting. Keratin contamination was suspected[c]. Even with samples carefully preparing in a HEPA-filtered laminar flow cabinet, contamination was still present. Mtu-2 and Mtu-3 had their mass measured successfully within 1 Da of the predicted value. For the other mutants, full-length DNA sequences were obtained of the mutant genes to ensure no other mutations had been inadvertantly introduced.

2.6.2 Secondary Structure

CD spectroscopy was used to ensure that the mutations had not introduced any major changes to the fold of the enzymes. Figure 2.9 shows that all of the mutants adopt similar secondary structure profile to the wild-type, implying similar folding.

2.6.3 Thermal Stability

A thermal melting curve was obtained by DSF for *Mtu*IPMS wild-type, as shown in Figure 2.10. The steepest point of the fluorescence curve was found where the second differential equalled zero (corresponding to the local maximum of the first differential). The temperature at this point was taken to be the melting

[c]Dr. Alexander Goroncy, Personal communication

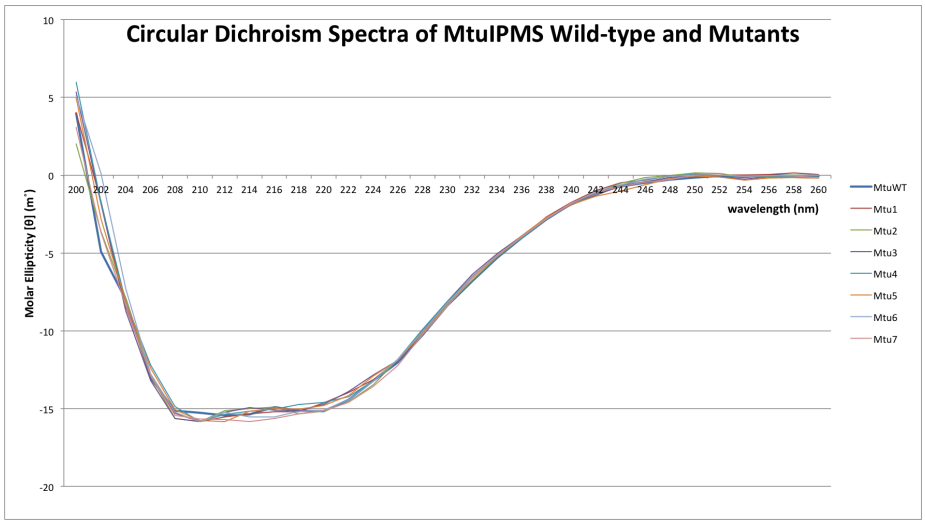


Figure 2.9: CD spectra of *Mtu*IPMS wild-type and mutant enzymes.

(unfolding) temperature of the protein. Readings were taken in triplicate and averaged. Figure 2.10 shows the fluorescence curve, readings from a control well with no protein, and the first differential. The second differential was omitted here for clarity.

As seen in Figure 2.4, DSF seems to be a relatively poor predictor of whether a given amino acid will inhibit *Mtu*IPMS mutants, at least under the conditions previously employed. To explore this area more fully, *Mtu*IPMS wild-type and several of the mutants from this study were probed using a range of amino acid concentrations. We also expanded the range of amino acids tested to include valine, so as to more fully explore this chemical space (see Figure 2.3).

*Mtu*IPMS wild-type and Mtu-1, 2, 3, and 5 were subjected to DSF to determine their melting point. As can be seen in Figure 2.11, these mutants had melting temperatures roughly 0.2°C to 1.2°C higher than the wild-type under than these conditions. This slight gain in stability was consistent with the range of melting temperatures observed in the previous study's mutants.

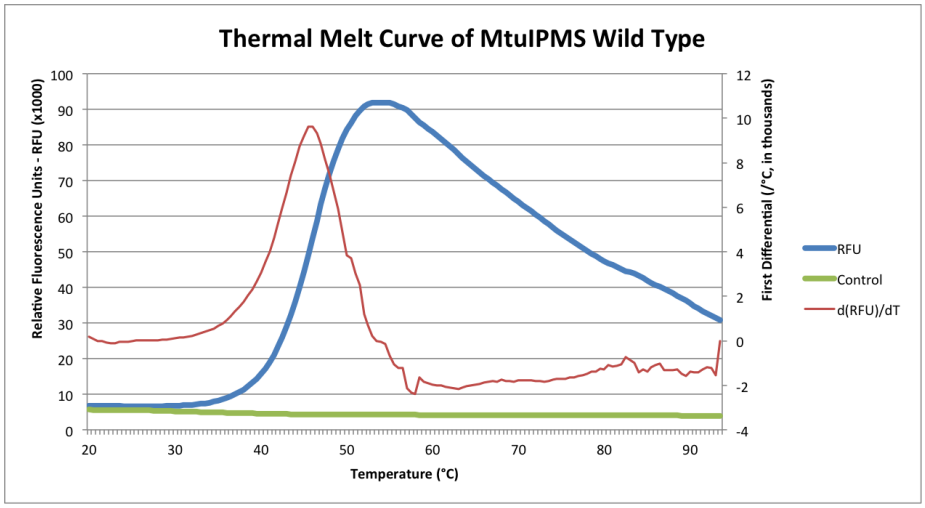


Figure 2.10: Thermal melt curve of *Mtu*IPMS wild-type. First differential curve and control wells shown.

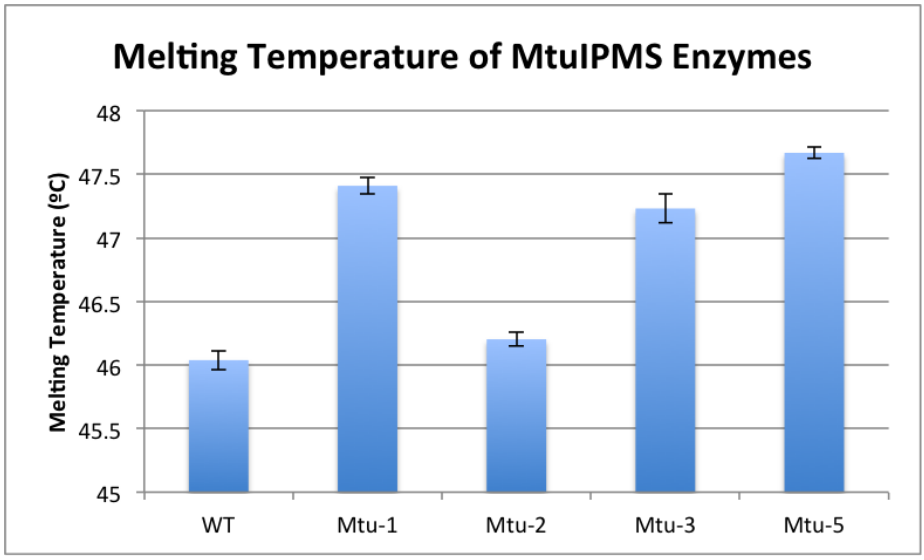


Figure 2.11: *Mtu*IPMS mutant enzyme melting temperatures

As well as the 5 mM concentrations of amino acids used in the previous study, a more physiologically realistic 0.5 mM concentration was also tested. The stabilisation of each enzyme with potential inhibitors can be seen in Figure 2.12.

Based on the stabilisation data, we could make several predictions about allosteric sensitivity, as displayed in Table 2.4.

Table 2.4: Predictions of allosteric sensitivity based on DSF data

	Mtu-				
Amino acid	**WT**	**1**	**2**	**3**	**5**
Leucine	Strong	—	Moderate	—	Weak
Isoleucine	—	Weakly Activating	Weak	Weak	—
Norvaline	Moderate	—	Weak	Weak	—
Valine	—	Weakly Activating	Weak	Weak	—

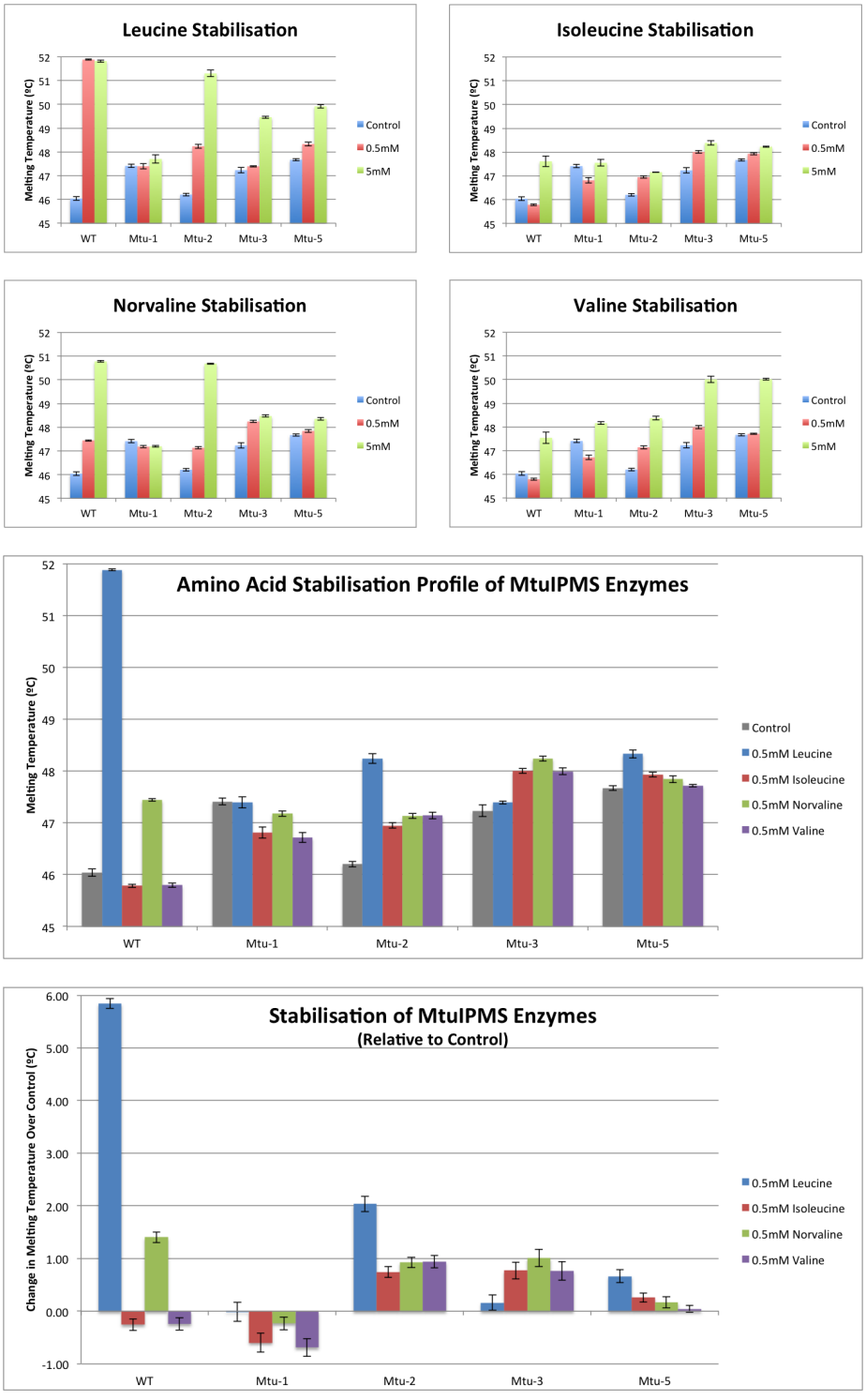


Figure 2.12: Stabilisation observed with *Mtu*IPMS enzymes in the presence of varying concentrations of amino acids.

2.7 Kinetic Characterisation

Due to time constraints K_m and k_{cat} values have not yet been measured for the mutants in this study. However, preliminary kinetic inhibition studies are presented here.

As can be seen from Figure 2.13 by adding in more mutations to our base A567V mutation, we can dramatically alter the allosteric sensitivity of *Mtu*IPMS. Particularly satisfactory is the observation that Mtu-4 (only two mutations away from the wild-type) is, at every concentration tested, significantly more inhibited by isoleucine than leucine. We have achieved one of the principle goals of this project: to switch the allosteric sensitivity of *Mtu*IPMS from leucine to isoleucine. Of further interest is Mtu-2 (again, two mutations away from the wild-type), which retains similar sensitivity to leucine as the wild-type, but is now just as sensitive to norvaline, and nearly as sensitive to isoleucine. We appear to have created a much more allosterically promiscuous enzyme. Another mutant worth noting is Mtu-6 (three mutations from the wild-type), which shows a similar inhibition pattern to Mtu-2, but is comparatively more sensitive to isoleucine than valine.

2.7.1 Correlation with DSF Predictions

Figure 2.14 shows that melting temperatures and inhibition rates are moderately well correlated, for both 0.5 mM and 5.0 mM concentrations. Data shown are for the mutants subjected to DSF in this study, and the mutants from the previous study. DSF was not carried out on several mutants (Mtu-4, -6, & -7) as despite the correlation, individual results can be seen to be highly unreliable in predicting allosteric activity. For instance, according to the figure, a gain of 2°C melting temperature could correspond to anywhere between 8% and 95% reduction in activity. Table 2.5 summarises the success of predictions made from

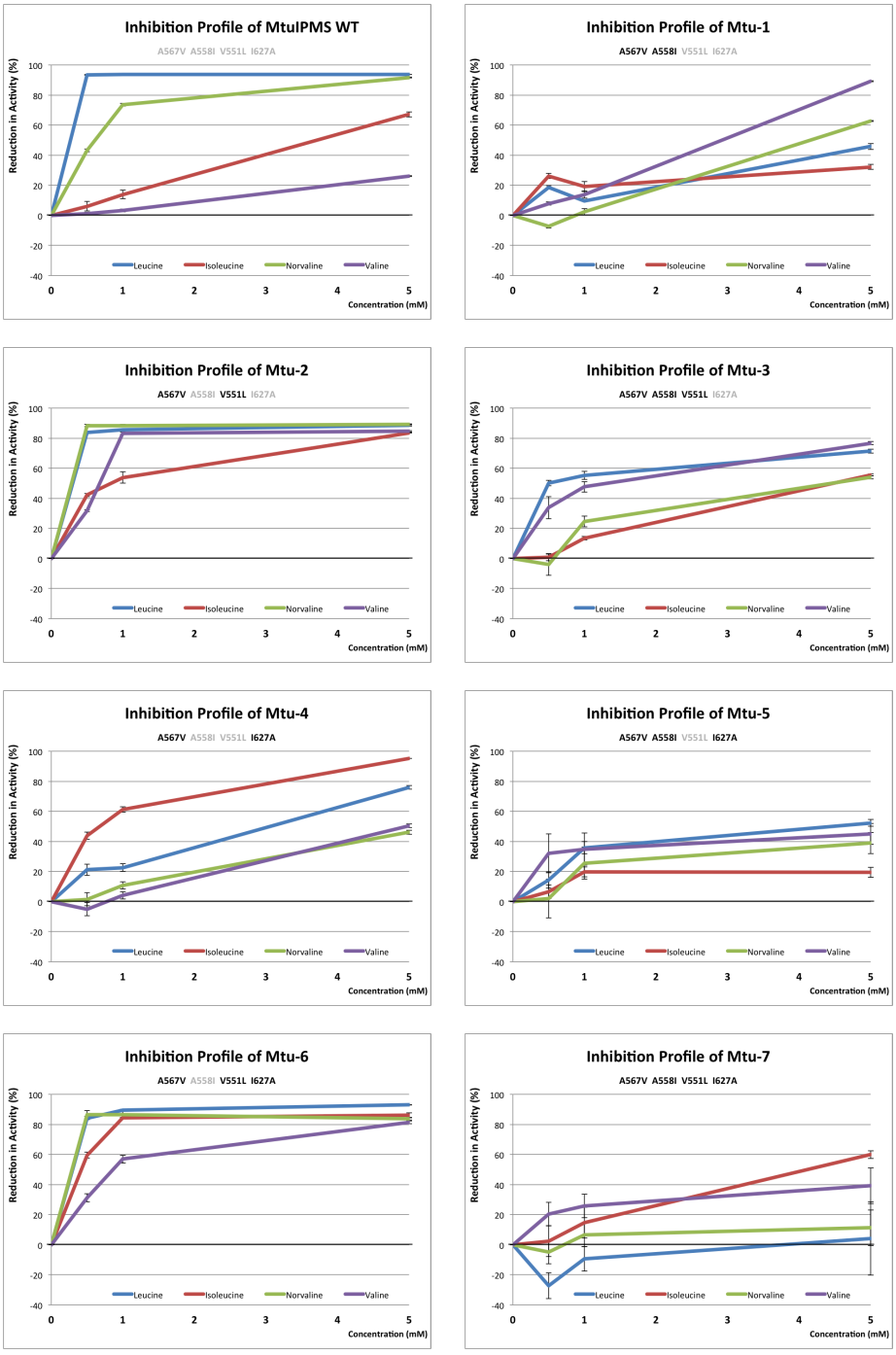


Figure 2.13: Inhibition of *Mtu*IPMS enzymes by different amino acids, at concentrations ranging from 0mM to 5mM. Activities were measured in 250μM α-KIV and 425μM AcCoA. The K_m values for the wild-type enzyme are 12 μM and 136 μM respectively. The concentration of AcCoA is not quite saturating, due to the high K_m for this substrate.

Table 2.5: Testing DSF predictions of allostery. DSF proved a reasonable, although not entirely reliable guide to allosteric activity.

	Mtu-				
Amino acid	**WT**	**1**	**2**	**3**	**5**
Leucine	✓✓	∼	✓	✗	✓✓
Isoleucine	∼	✗	✓	✓✓	✓✓
Norvaline	✓	✗	✓	✓✓	✓
Valine	✗	✗	✓	✓	✓

✓✓: DSF correctly predicted the degree of inhibition.
✓: predicted whether or not there was inhibition.
∼: no inhibition predicted, only weak inhibition observed.
✗: activation or no inhibition predicted, moderately strong inhibition observed.

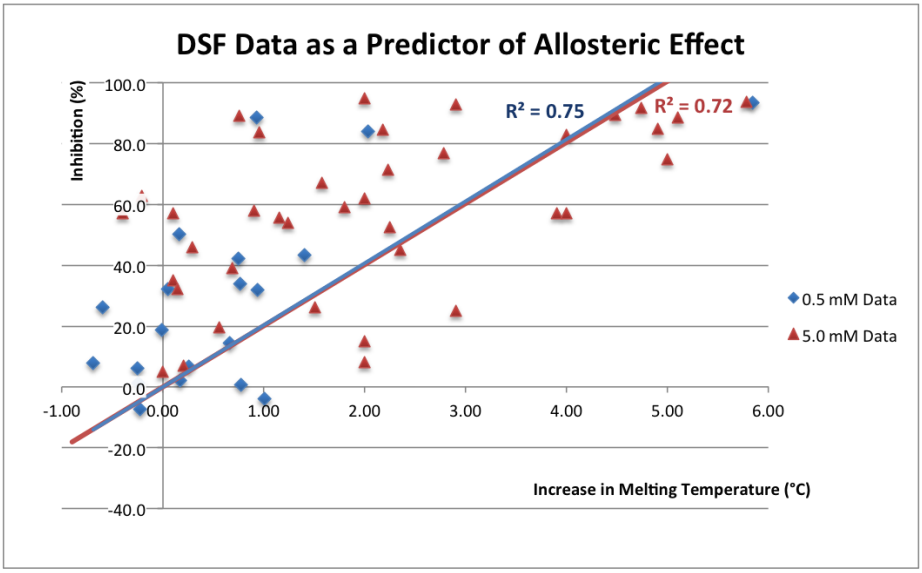


Figure 2.14: Using the data from several mutants in this study, and the previous study, DSF does show moderate predictive power for inhibition.

the DSF results in this study. As can be seen from the table, the predictive success of DSF was rather mixed, so we chose not to characterise the remaining mutants by DSF, opting to focus on kinetic studies.

2.8 Summary

Seven new variants of *Mtu*IPMS were created, purified, and characterised by DSF and kinetic inhibition studies. While DSF is often cited as a good first-pass approximation for protein-ligand binding, our study suggests that, at least for *Mtu*IPMS, this isn't a reliable screening tool. While there is a good overall predictive value for DSF results, if we had chosen which mutants to kinetically characterise based purely on DSF, we would have missed many interesting allosteric interactions.

The goal of this project was to switch the allosteric specificity from leucine to isoleucine, and in the case of Mtu-4, this was a success. At every concentration tested, isoleucine was a significantly better inhibitor than leucine. Based on the crystal structure, Ile627 was the major barrier to isoleucine binding, as it sterically blocked the β-methyl group of its sidechain. Ile627 also seems to provide favourable hydrophobic interactions with the β-carbon of leucine and norvaline. Ala567 seems to stabilise the terminal isopropyl group of leucine, while valine in this position does a better job of stabilising isoleucine's terminal methyl carbon. The valine mutants would also restrict the space available for leucine's isopropyl terminus. Mutating Ile627 to alanine was not enough to allow isoleucine to bind by itself, but with A567V it seems we have better provided the hydrophobic interactions needed to stabilise isoleucine, while tuning the steric factors to disfavour leucine binding.

Mtu-2 and Mtu-6 have both become rather allosterically promiscuous compared to the wild-type with the addition of V551L. This is the most conservative

mutation explored here, both structurally and sequence-wise: Figure 2.2 shows that these two residues are commonly interchanged among α-IPMS enzymes. The extra methylene in leucine gives the sidechain more flexibility, which could explain why it allows the allosteric pocket to accommodate a wider range of inhibitors.

A558I seems to be an awkward mutation in *Mtu*IPMS. It consistently lowers the allosteric sensitivity for all molecules whenever it is present, and Clarke showed it lowers the melting point of wild-type *Mtu*IPMS. PyMOL modelling suggests the least sterically hindered conformation it can adopt places the ethyl group of its sidechain in a polar environment, as can be seen in Figure 2.15. We would predict that this would both lower the stability (explaining the lowered melting point) and project a methyl group into the allosteric pocket, reducing the ability of our inhibitors to bind.

We expected that Mtu-7, carrying all of the mutations identified from CMS, should be the best at binding isoleucine: but in fact it shows a fairly poor inhibitory profile. This could be due to the destabilising effect of A558I, and indeed Mtu-6 (which carries the other three mutations) does have the strongest response to isoleucine. However, as discussed above, V551L seems to allow greater promiscuity, so even though Mtu-4 shows slightly less sensitivity to isoleucine than Mtu-6 does, it is more *selective* for isoleucine.

Exploring the role of A558I by crystallography would be worthwhile, to confirm our speculations as to how it is affecting the allosteric site. It would be interesting to crystallise some of the other enzymes in this study with various inhibitors bound, to study the exact nature of the changes we have made. Now that we have caused some interesting changes to the allosteric properties, it would be worthwhile to construct mutants without A567V, to see whether this mutation is, in fact, critical for isoleucine sensitivity.

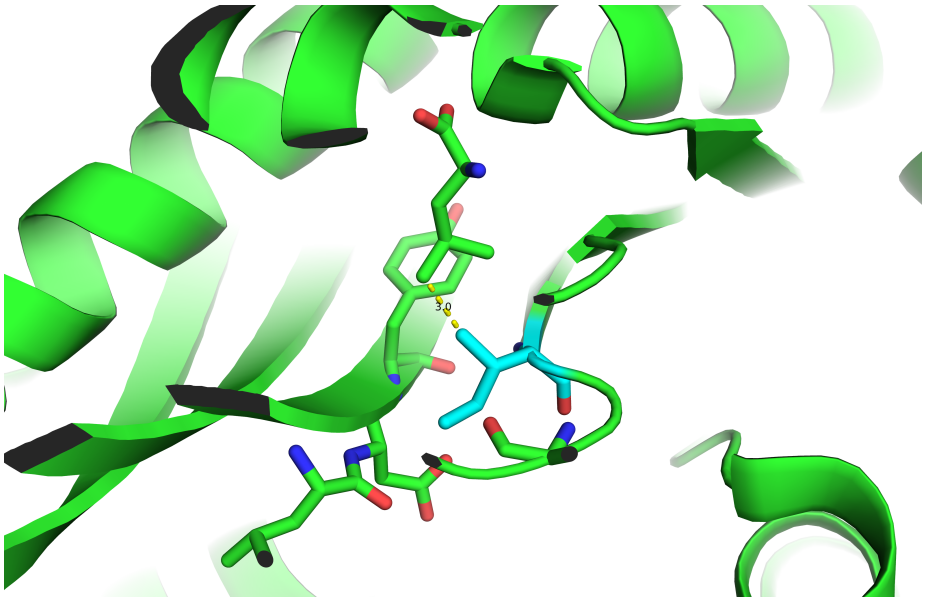


Figure 2.15: A558I reduces inhibition and stability of the enzyme. The ethyl group is forced into a polar pocket of carbonyl and hydroxy groups, while the upwards-protruting methyl group sterically hinders inhibitor binding. PyMOL model based on PDB ID: 3FIG.

Chapter 3

Probing the Allosteric Mechanism of *Neisseria meningitidis* α-IPMS via Molecular Dynamics Predictions

3.1 Overview

In the previous chapter we used insights gleaned from the crystal structure of *Mtu*IPMS to successfully make rational mutations, altering its allosteric specificity. *Neisseria meningitidis* α-IPMS (*Nme*IPMS) is a smaller protein ($\sim$56 kDa vs $\sim$70 kDa) and has fewer decorations (for instance, *Mtu*IPMS has

variable numbers of tandem repeats[22] and an N-terminal extension[9] absent in *Nme*IPMS). Due to these considerations, and as there is no complete crystal structure for *Nme*IPMS, Dr. Wanting Jiao has constructed a homology model for molecular dynamics (MD) simulations. Both enzymes are feedback inhibited by L-leucine through altered molecular dynamics rather than alterations to the tertiary or quaternary structure.[9,24,25,49] This work tests *in vitro* the predictions made by MD simulations *in silico*, in order to shed light on the mechanism by which the signal is conducted from the allosteric pocket to the active site.

The asymmetric, domain-swapped nature of the α-IPMS dimer, first described in *Mtu*IPMS (PDB ID: 1SR9 etc.) is preserved in the homology model of *Nme*IPMS. The catalytic barrel of one chain is capped by subdomain I of the other chain, and positioned adjacent to subdomain II and the regulatory domain of that second chain. The active site on one side of the dimer is relatively open, while subdomain II and the regulatory domain sit much closer to the catalytic site on the other side. Figure 3.1 shows the overall structure of the model and Figure 3.2 shows the domain-swapped nature of the dimer. The asymmetry of the dimer is curious: what, if any, functional purpose does the asymmetry serve? It is unclear whether the active site is catalytically active in both the "open" and "closed" conformations, or whether only one catalytic site per dimer is active. In which case, is the other chain merely a complicated structural prop? Or perhaps the protein undergoes a pendulum-like rearrangement, with the regulatory domain flopping back and forth, allowing both sites to catalyse reactions. This may play a role in allostery, or perhaps in substrate binding and product release. Similar questions are considered in several systems in a review by Nagradova.[50] So far, solution-phase studies and mutagenesis on *Mtu*IPMS[51] have revealed few clues, and initial MD simulations indicate a prohibitive energetic barrier to any such rearrangements[a]. However,

[a] Personal communications, Dr. Wanting Jiao

that leaves unexplained the reason for the asymmetry. Domain movement in α-IPMS enzymes will be briefly examined in the experiments here, which promises some interesting preliminary results.

Several residues of interest were identified by Dr. Jiao for experimentation. These were mutated and the corresponding proteins purified for analysis. Mass spectra were obtained for most of the enzymes and were in good agreement with the predicted mass. The secondary structure was investigated by circular dichroism (CD) spectroscopy, and no evidence of misfolding was found. Preliminary kinetics testing has been carried out and the results warrant further investigation.

3.2 Residues Selected for Analysis

The following residues were identified by Dr. Jiao for experimental analysis. This information is summarised in Table 3.1.

3.2.1 Tyr313

Studies have identified Tyr410 as a key residue mediating allostery in *Mtu*IPMS.[24,29] Multiple sequence alignments (MSA) show this is strongly conserved, and corresponds to Tyr313 in *Nme*IPMS (Figure 3.3). In *Mtu*IPMS, Koon *et al.*[9] identified Tyr410 as being responsible for orientating the catalyticaly important His379 residue. Substituting phenylalanine for tyrosine at this position has been shown to completely eliminate inhibition by leucine.[29] While leucine binding seems to be unaffected, H/D exchange experiments[24] show residues near the active site adopt similar dynamics (and catalytic rate) to the inhibited enzyme, regardless of whether leucine is present. As this tyrosine residue seems to be so well conserved, we would expect the same effect should be present in *Nme*IPMS.

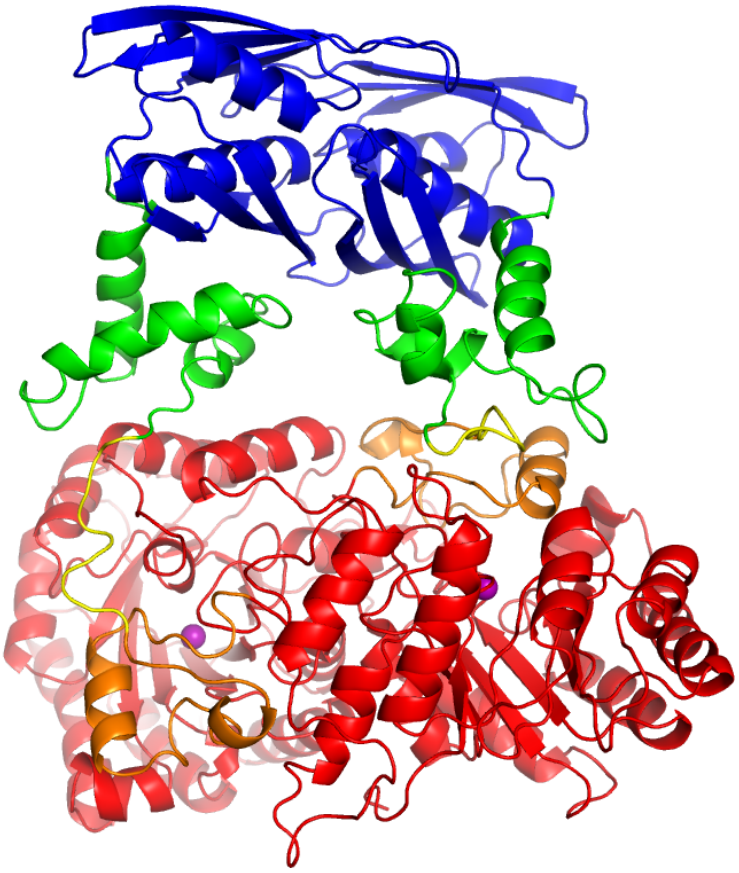


Figure 3.1: Rendering of the *Nme*IPMS homology model. Catalytic domain shown in red, subdomain I in orange, linker in yellow, subdomain II in green, and the regulatory domain in blue. The asymmetrical arrangement of the homodimer is clearly visible. On the left side, as shown here, the end of the catalytic barrel (Zn shown in pink to denote the active site) is relatively uncovered, whereas on the right, it is somewhat covered by subdomain II and the regulatory domain. We will refer to these as the "open" and "closed" conformations, respectively.

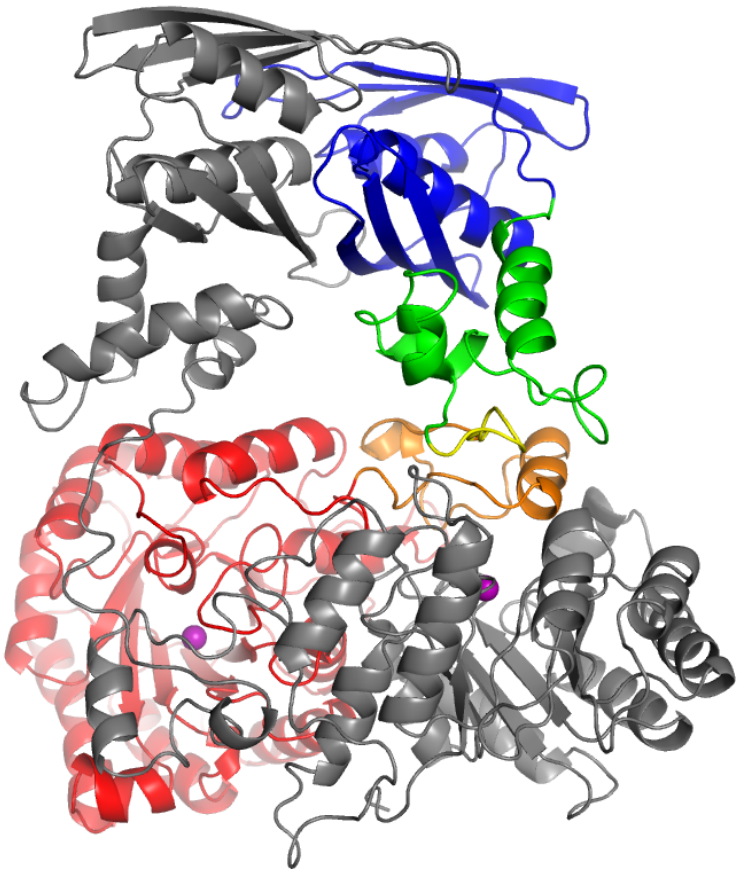


Figure 3.2: *Nme*IPMS homology model, chain B coloured grey. The domain-swapped nature of the enzyme can be clearly seen: subdomains I, II, and the regulatory domain of chain B are all on the same side as the catalytic barrel of chain A, and vice versa.

As a key catalytic residue, substitution with alanine ought to drastically lower the catalytic rate.

3.2.2 Arg32 & Asp375

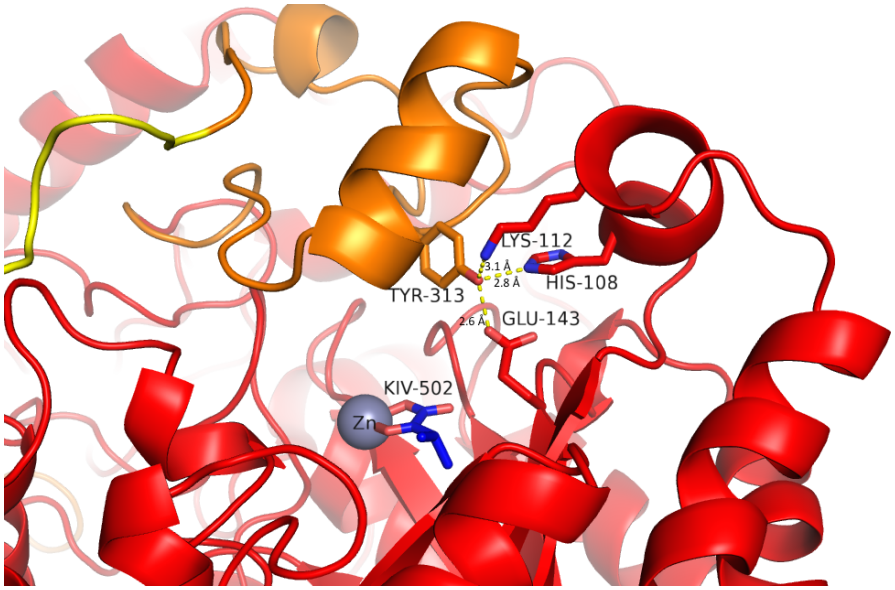
On one side of the asymmetric dimer, Arg32 and Asp375 appear to form a salt bridge (Figure 3.4), stabilising the asymmetric conformation we see from the *Mtu*IPMS crystal structures, with one chain in the "closed" conformation and the other "open". MD simulations[b] indicate that when leucine binds to the regulatory domain, the dimer forms a more rigid, symmetrical conformation. By mutating either Arg32 or Asp375 to cysteine, we should be able to disrupt the salt-bridge and see whether it mimics the effect of leucine binding, or uncouples leucine binding from catalysis. It may even increase the catalytic rate, if rearrangement between asymmetric conformations is required for catalysis.

If both residues are mutated to cysteine, we would predict that a disulfide bridge would form between them, preventing any pendulum-like rearrangement of the "open" and "closed" conformations, and disallowing any intermediate symmetrical states. If the enzyme does indeed undergo rearrangement as part of allostery or catalysis, fixing it in one conformation should drastically alter its kinetic behaviour. Previous work by Dr. Frances Huisman [52] achieved a similar disulfide-forming mutant in *Mtu*IPMS, and found a substantial reduction in activity. However, it was shown that one of the mutations by itself was sufficient to provoke this reduction in activity, and so it was not clear whether the disulfide bridge had any effect. In this work, both single cysteine mutations in *Nme*IPMS (R32C and D375C) are isolated and assayed. If these mutants do not alter catalytis significantly, it would be sensible to construct the double-mutant and see whether it forms a disulfide bridge. If it does, any change in catalysis or regulation must be due to the restriction of domain movements. This would

[b] Dr. Wanting Jiao, unpublished work

	Nme Y313
Mycobacterium_tuberculosis	--DCDVDDMLWQVPYLPIDPRDVGRTY-
Neisseria_meningitidis	----------ETYEIMSAESVGWAT-
Escherichia_coli	----------ENYEIMTPESIGLNQ-
Saccharomyces_cerevisia	------GETQWRIPYLPLDPKDIGRDY-
Liseteria_monocytogenes	----------DTYEIITPALVGVDK-
Haemophilus_influenzae	----------NTYEILSPETIGLKK-
Salmonella_enterica	----------ENYEIMTPESIGLNQ-
Streptomyces_coelicolor	--GVTVDDIEWAVPYLPIDPKDVGRSY-
Aradopsis_thaliana	----------GTYEIICPEEIGLERS
Buchnera_aphidicola	----------KNYEIMEPSSIGLKE-
Bordetella_bronchiseptica	------ADAPWEVPYLPIDPADLGRSY-
Bifidobacterium_longum	--GADLDSFVWLVPYLPIDPKDIGRTY-
Corynebacterium_glutamicum	VSWEQLRDTEWEVPYLPIDPKDVGRDY-
Nocardia_farcinica	--NADVSDIVWEVPYLPIDPKDVGRTY-
Gordonia_bronchialis	--DRDVDDILWQVPYLPIDPKDVGRNY-

(a) Tyr410 is seen to be highly conserved in this MSA constructed using Jalview/ClustalO.[45,46]



(b) Active site of *Nme*IPMS from homology model. Y313 is predicted to have polar contacts with several nearby residues. Catalytic zinc ion and α-KIV are shown to denote the active site. Catalytic domain is coloured in red, subdomain I in orange, and the linker in yellow.

Figure 3.3: *Mtu*IPMS Y410 is a strongly conserved residue, and has been implicated in mediating the allosteric signal. The corresponding residue in *Nme*IPMS is Y313.

provide insight into the role of any pendulum-like rearrangements in allostery or catalysis in α-IPMS enzymes.

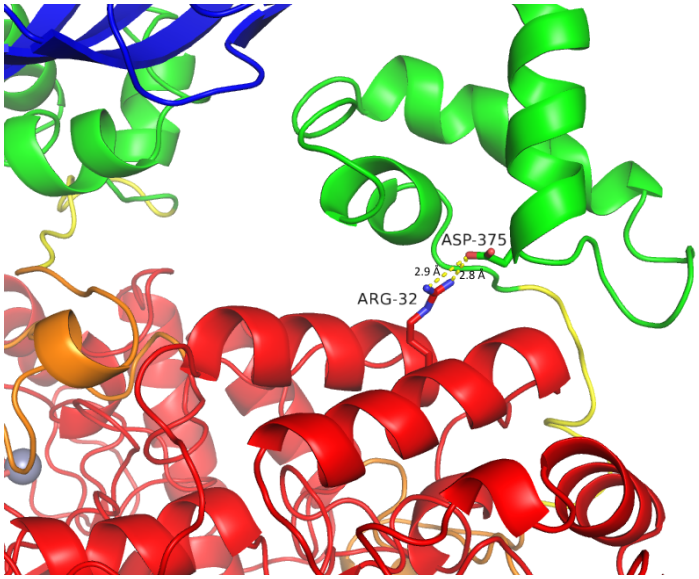


Figure 3.4: Homology model of *Nme*IPMS. By disrupting the salt-bridge formed by Arg32 and Asp375, or locking it into a disulfide bridge, we hope to learn more about the role of flexibility between the regulatory and catalytic domains in α-IPMS enzymes.

3.2.3 Lys332

In the leucine-bound MD simulation, subdomain I is lifted slightly out of the catalytic site: this may be how the catalytic rate is reduced in the inhibited enzyme. Lys332 is part of the linker between subdomains I and II, and forms polar interactions with Glu18 and Asp56 (see Figure 3.5) on the catalytic barrel. This seems to restrict the ability of subdomain I to move away from the catalytic barrel. Mutating Lys332 to alanine may therefore disrupt catalysis or inhibition.

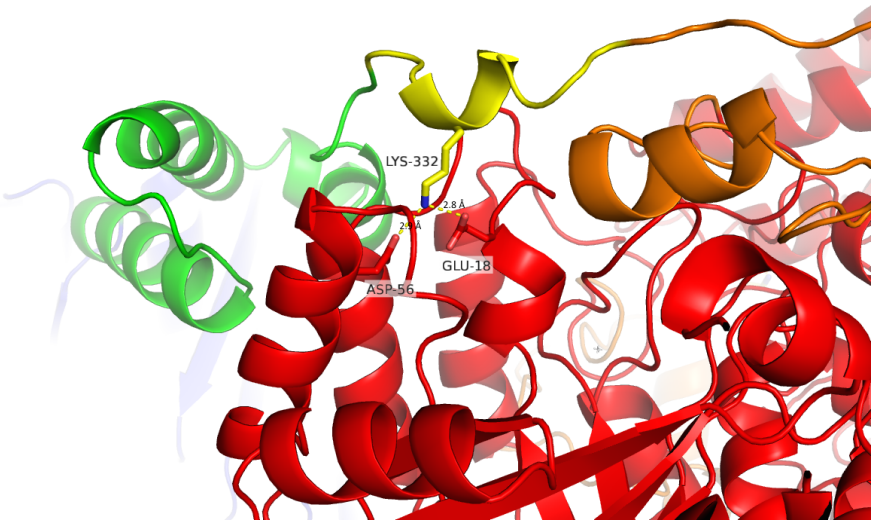


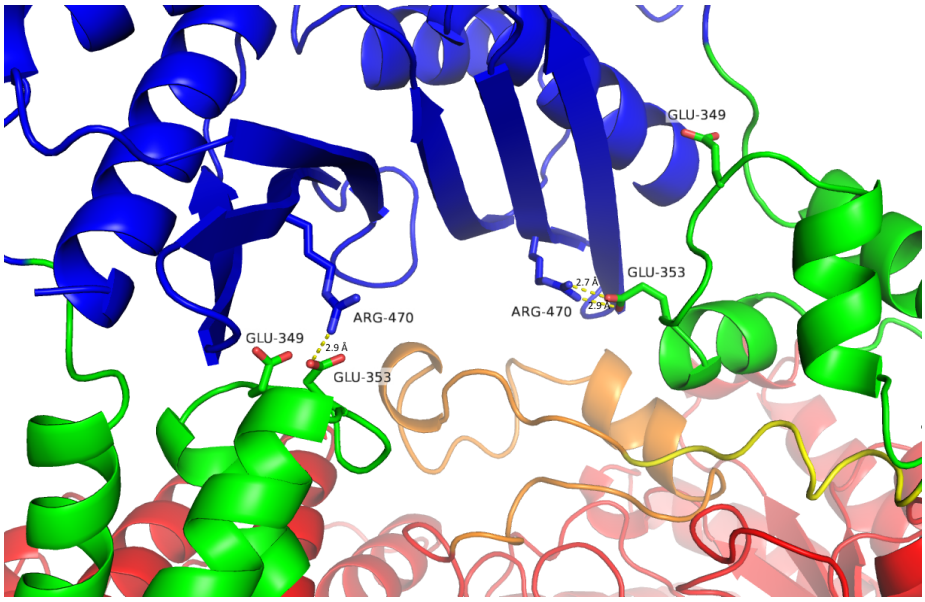
Figure 3.5: Homology model of *Nme*IPMS. Lys332 may play a role in mediating allostery, restricting movement of subdomain I during leucine binding.

3.2.4 Arg470

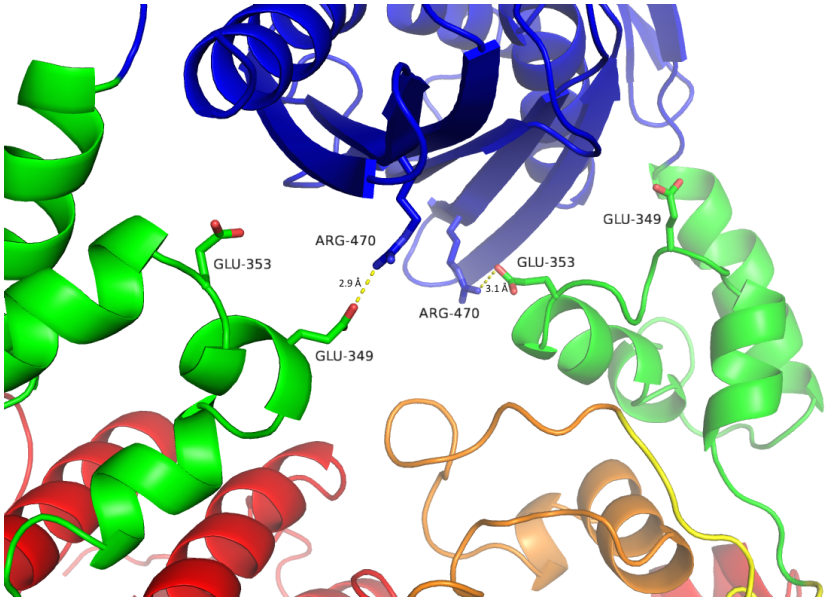
When leucine binds the allosteric domain, Arg470 on one chain loses a hydrogen bond to Glu353, and forms one with Glu349 (Figure 3.6). Previous work by Tyler Clarke[1] showed that mutating Glu353 alone is not enough to disrupt inhibition, but Arg470 looks more promising as a mutagenesis target. By disrupting this shifting hydrogen bond network we may be able to disrupt inhibition.

3.3 Generation of Mutants

Site-directed mutagenesis (SDM) proceeded as per the previous chapter. A commercially available plasmid-extraction kit was used to isolate template DNA from a small overnight growth of an *E. coli* plasmid propagation strain con-



(a) No leucine.



(b) Leucine bound.

Figure 3.6: Molecular dynamics simulation structures kindly supplied by Dr. Wanting Jiao. One one chain Arg470 changes hydrogen bonds between Glu353 and Glu349 when leucine binds.

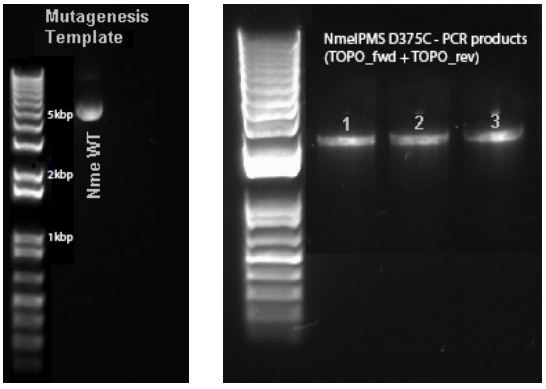
Table 3.1: *Nme*IPMS mutants identified for testing.

Residue	Mutation	Rationale
Y313	F / A	Residue corresponding to Tyr410 in *Mtu*IPMS, found to uncouple leucine binding from inhibition. We would predict the same effect in *Nme*IPMS.
R32 / D375	C	Form a salt-bridge between the catalyic barrel and subdomain II when the enzyme is in its "open" conformation. Disrupting this bond may interupt allosteric signal conduction. Residues were mutated to cysteine in preparation for the construction of a double-mutant which may be able to form a disulfide bond. This would tell us whether alternating between "open" and "closed" conformations is important for catalysis and/or regulation.
K332	A	May hold subdomain I clear of the catalytic site upon leucine binding, or prevent it from moving too far away. Destroying this interaction may remove leucine sensitivity or disrupt catalysis.
R470	A	Forms salt-bridges with two nearby residues: Glu353 when no leucine is present, and Glu349 upon leucine binding. Removing these contacts may remove leucine-sensitivity.

taining the plasmid pFH01. This plasmid is a pET-151 vector carrying the *Nme*IPMS gene, a T7 promotor to induce overexpression with IPTG, N-terminal $(His)_6$-tag, TEV-protease recognition site for tag removal, and a gene encoding for ampicillin-resistance as a selection marker. Plasmid is shown along with a ladder for size in Figure 3.7a).

The template was subjected to SDM to introduce the desired mutations, and the SDM reaction product was digested using DpnI restriction enzyme to remove remaining template DNA. The DpnI digest was transformed into chemically competent *E. coli* cells which were grown overnight, and several colonies picked for plasmid extraction. Plasmids were subjected to PCR with sequencing primers, and both template and PCR product separated on agarose to check integrity, similar to that seen in Figure 3.7b. Plasmids were then submitted for DNA sequencing to confirm the mutation and gene integrity. A

sequence-confirmed plasmid was then used to innoculate a chemically competent *E. coli* growth strain, which was stored as a -80°C glycerol stock, until required for growth and protein extraction.



(a) *Nme*IPMS wild-type in pFH01.

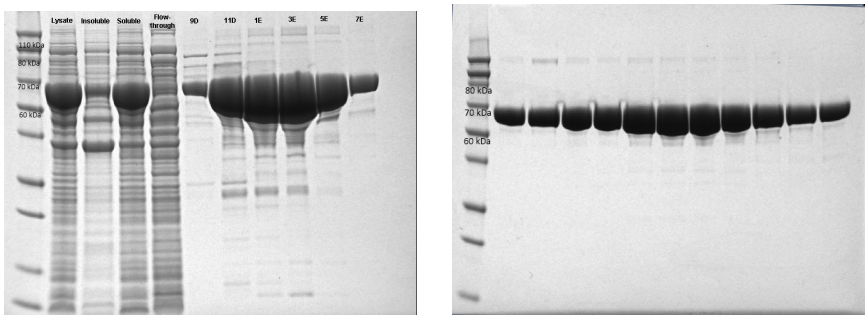
(b) PCR products of *Nme*IPMS mutant plasmids checked before sequencing.

Figure 3.7: DNA in the form of wild-type and mutant plasmids, and their PCR products, were separated on agarose gel electrophoresis to check plasmid integrity.

3.4 Protein Expression and Purification

Wild-type and mutant *Nme*IPMS enzymes were mostly grown and expressed as per the *Mtu*IPMS enzymes from the previous chapter: full details in 5.4. However, unacceptable loss of protein often occurred with the *Nme*IPMS variants while attempting to cleave the polyhistidine tag. Presumably the natural N-terminal extension present in *Mtu*IPMS increases the accessibility of the protease recognition site, allowing much easier removal of the tag. Due to this problem, the wild-type *Nme*IPMS was purified in both tagged and untagged form, while the mutants were purified with tags intact to allow for easier screening. It was found that the N-terminal tag did have a moderate

effect on catalysis (Table 3.3) but no discernable effect on inhibition (data not shown).



(a) HisTrap (b) Size-Exclusion Chromatography

Figure 3.8: *Nme*IPMS-Y313F purification visualised by SDS-PAGE.

3.4.1 Protein Extraction

Fresh or thawed cell pellets were resuspended and lysed by sonication on ice with Benzonase® Nuclease (Sigma) to break down nucleic acids. Lysate was clarified by centrifugation and filtration, and the insoluble pellet resuspended in 8M urea for electrophoresis. The lysate and soluble fractions in Figure 3.8a show clear overexpression of the tagged *Nme*IPMS enzyme at 59 kDa, but the insoluble fraction still contains a certain amount of protein, suggesting incomplete lysis.

3.4.2 Immobilised Metal Affinity Chromatography (IMAC)

The clarified soluble fraction was loaded onto a HisTrap (GE Healthcare) IMAC column. These columns coordinate Ni^{2+} ions, but leave several coordination sites free. These sites are bound by imidazole moieties on polyhistidine tags (His-tags), causing highly specific binding of tagged proteins to the column. An imidazole gradient is used to elute the protein, which is monitored by absorption at 280nm. The protein elutes reasonably pure, as can be seen in the fractions

gelled in Figure 3.8a. Fractions containing the bulk of the protein were pooled. Samples to have their His-tag cleaved were treated on a desalting column to remove imidazole, before the following step (3.4.3). Proteins purified in their tagged state were simply concentrated and polished on the SEC column (3.4.4).

3.4.3 TEV Protease Treatment

To remove the polyhistidine tag, the sample was incubated with TEV protease, which cleaves the amide bond at a recognition sequence designed into the short linker between the tag and the beginning of the *Nme*IPMS sequence. This sequence is described in Figure 3.9.

MHHHHHHGKPIPNPLLGLDSTENLYFQGIPFT

Figure 3.9: The hexahistidine section (red) binds to the IMAC column, followed by a linker region (orange), and a TEV protease recognition site (underlined). The GIPFT residues (blue) after the recognition site remain attached to the protein.

Approximately 1mg TEV protease was added per 10-100mg tagged protein, along with DTT and EDTA. The sample was incubated for approximately 1 hour on the bench to allow the reaction to proceed, before being incubated overnight at 4°C. The sample was desalted again to remove the DTT and EDTA, before being subjected to a second IMAC step. This bound the His-tagged TEV protease enzyme and the cleaved His-tag, allowing the cleaved protein to elute.

3.4.4 Size-Exclusion Chromatography (SEC)

A final polishing step was carried out via SEC. The enzyme eluted as a single peak, and fractions were examined by SDS-PAGE (Figure 3.8b). The mass inferred from the gel (≈ 65 kDa) matches the predicted protein size. All fractions visually judged to be >95% pure were pooled and concentrated to $>$ 1mg/mL, and flash-frozen in liquid nitrogen before being stored at -80°C.

3.5 Physical characterisation

3.5.1 Mass Spectrometry

Mass spectrometry measurements were in good agreement with all predicted values, as shown in Table 3.2. As with *Mtu*IPMS, some proteins were not able to be measured (i.e. wild-type: tagged and untagged) but full sequencing was obtained to ensure correct sequence and gene integrity.

Table 3.2: Mass spectrometry results for *Nme*IPMS mutants. ND (not determined) means the mass was not able to be measured.

Mutant	Measured Mass (Da)	Predicted Mass (Da)
Wild-type	ND	59,179
Wild-type (tag removed)	ND	56,029
Y313A	59,087	59,088
Y313F	59,162	59,164
D375C	59,166	59,168
R32C	59,124	59,127
K332A	59,122	59,123
R470A	59,095	59,095

3.5.2 Secondary Structure

CD spectroscopy was carried out on all samples to ensure the mutations had not result in any misfolding. As can be seen in Figure 3.10 the mutant proteins all adopted a similar fold to the wild-type.

3.6 Preliminary Kinetic Characterisation

Due to time constraints and limited substrate availability, only preliminary kinetic results were obtained. Table 3.3 shows the enzymatic activity of each

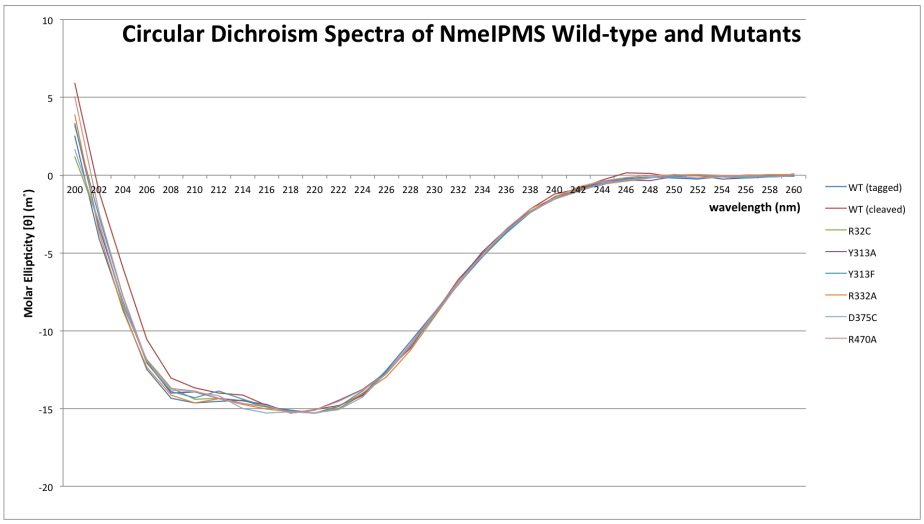
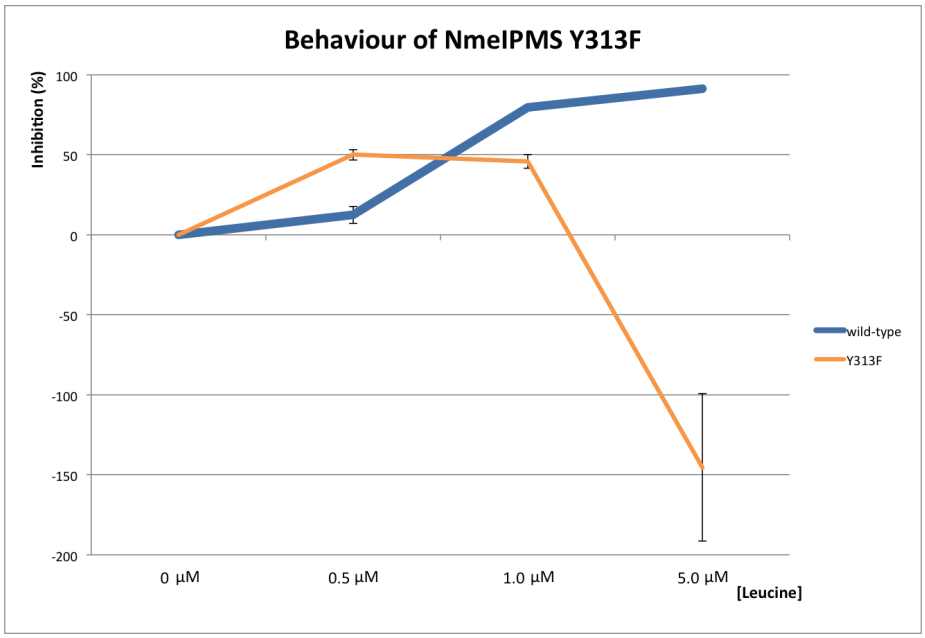


Figure 3.10: CD spectra of *Nme*IPMS wild-type and mutant enzymes.

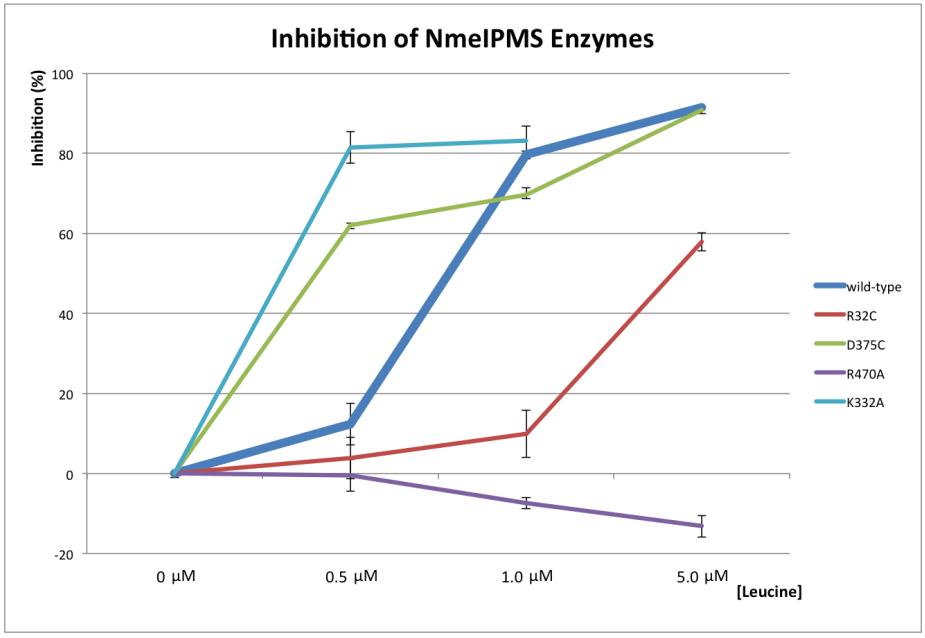
enzyme studied in the presence of 250 μM α-KIV and AcCoA, and the results of inhibitions studies are presented in Figure 3.11.

The untagged wild-type enzyme shows approximately six-fold increased activity over the tagged wild-type, indicating the tag does exercise a moderate effect on catalysis. The rate observed here is consistent with the previously reported[32] k_{cat} of $13.0 \pm 0.3\, s^{-1}$. Further investigations should therefore work on the untagged protein. All other comparisons in this section will use the tagged wild-type and mutants to control for any effect of the tag.

As we might expect, Tyr313 seems to be rather important for catalysis. Substituting it with alanine lowers the catalytic rate by a factor of nearly 700. The more conservative phenylalanine substitution still resulted in a ~80-fold reduction in activity, comparable with the ~30-fold reduction de Carvalho *et al.*[29] found with *Mtu*IPMS Y410F. However, the cited study found that in *Mtu*IPMS, the Y410F mutant was insensitive to leucine up to 500 μM, whereas our *Nme*IPMS Y313F mutant only showed ~50% activity at 0.5-1.0 μM



(a) Inhibition in *Nme*IPMS wild-type and Y313F.



(b) Inhibition in *Nme*IPMS wild-type and several mutants

Figure 3.11: Preliminary inhibition studies of *Nme*IPMS variants. Negative values indicate activation rather than inhibition.

Table 3.3: Summary of activity of *Nme*IPMS enzymes. The "change" column refers to an n-fold increase(+) or decrease(–) in activity over the tagged wild-type enzyme. Activities were measured under saturating conditions: 250 μM α-KIV and AcCoA. The K_m values for the wild-type enzyme are 30 μM and 35 μM respectively.

*Nme*IPMS-	Activity (/s)	Change
wild-type (untagged)	8.9 ± 0.2	—
wild-type	1.530 ± 0.002	—
R32C	3.70 ± 0.08	+2.42
D375C	3.4110 ± 0.0005	+2.229
R470A	4.120 ± 0.004	+2.67
K332A	0.024 ± 0.002	–63
Y313A	0.002 ± 0.002	–700
Y313F	0.020 ± 0.005	–80

concentrations of leucine, and greatly increased activity with 5.0 μM leucine (Figure 3.11a). Further studies are required to investigate this unusual result.

Arg32 and Asp375 were both mutated to cysteine, and unlike the corresponding mutations in *Mtu*IPMS, neither of these mutants resulted in lowered catalytic activity. R32C did substantially lower leucine sensitivity, indicating this salt-bridge is important in transmitting the allosteric signal. D375C showed an increase in leucine sensitivity at lower concentrations, and slightly decreased sensitivity at higher concentrations. Given the relatively similar size of the aspartate and cysteine sidechains, it is perhaps unsurprising that cysteine can adopt a similar role in this position: presumably it is acting as a hydrogen-bond acceptor from Arg32, propagating the allosteric signal via this bond just as aspartate would with via a salt-bridge. In any case, we have shown that polar interactions here are important for allostery: in future work, the double-mutant may show whether catalysis in interrupted by locking the domains with a disulfide bridge, perhaps shedding light on the question of domain movement in α-IPMS enzymes.

Mutating Lys332 was predicted to reduce leucine sensitivity, but the main effect seems to be that it has reduced the catalytic capacity of the enzyme ∼60-fold. Due to the low catalitic activity, no firm conclusions can be drawn as to whether allostery has been affected. As this residue looks to be important for keeping subdomain I in position for catalysis when leucine-binding lifts it from the active site, it is understandable that catalysis is compromised, and more sensitive to any disrupting influence (such as leucine binding).

As predicted, mutating Arg470 successfully removed all leucine inhibition, without drastically altering catalytic activity, proving the importance of the hydrogen-bonds predicted by MD simulations.

3.7 Summary

While the mutants described above have yet to be fully characterised, the initial results provide some rather enticing insights. The bizarre interaction of Y313F with leucine needs to be more fully explored. As well as a full kinetic characterisation, thermal stability ought to be examined with varying concentrations of leucine. It is possible that the decreased activity we see with Y313F is caused by increased structural flexibility associated with the mutation: Frantom *et al.*[24] found several regions of *Mtu*IPMS to have increased solvent exposure due to the corresponding Y410F mutation. In that study, and in previous work in this group[1,52] leucine binding appears to stabilise the enzyme, which may explain why higher concentrations of leucine restored some catalytic ability.

Predictions regarding the properties of two mutants (K332A and R470A) based on MD simulations were tested, and preliminary results suggest the simulations accurately reflect motions tested *in vitro*. Allosteric inhibition was completely destroyed in K332A, with no decrease in catalytic activity. This

indicates the predicted salt-bridges are essential for relaying the allosteric signal away from the regulatory domain. Arg470 appears to be a crucial residue for maintaining catalytic activity, and a more thorough exploration of its kinetics ought to shed more light on whether it also plays a role in allostery.

Unlike the cysteine mutants previously studied in *Mtu*IPMS, the *Nme*IPMS mutants (R32C and D375C) showed no evidence (at this preliminary stage) that the mutations harm catalysis. Interrupting the predicted salt-bridge does appear to decrease leucine sensitivity, although due to cysteine's ability to form hydrogen bonds, future work should confirm this by mutating to alanine. Another area to explore is to form a cysteine double-mutant. We predict it should form a disulfide bridge, and by determining effect this has on catalysis, we will hopefully shed some light on the questions surrounding domain movement and flexibility in α-IPMS enzymes.

Chapter 4

Discussion

The aim of this work was to probe the selectivity and mechanism of allosteric regulation in α-IPMS enzymes. We successfully altered the allosteric specificity of MtuIPMS based on insights gained from sequence homology and crystal structures of MtuIPMS (inhibited by leucine) and LiCMS (inhibited by isoleucine) enzymes. We also tested predictions from molecular dynamics (MD) simulations for the roles of several residues in *Nme*IPMS in conducting the allosteric signal from the regulatory domain to the catalytic site, a distance of ~ 50Å.

The approach followed was to introduce the desired mutations to *Mtu*IPMS and *Nme*IPMS genes using site-directed mutagenesis, and extracting and purifying the resultant enzymes using immobilised metal affinity and size-exclusion chromatography. These mutant enzymes were subjected to preliminary physical and kinetic characterisation to assess properties such as thermal stability and inhibition by various branched-chain amino acids.

One principle goal of this study was to tune the allosteric specificity of MtuIPMS from leucine towards isoleucine. Zhang *et al.*[15] had previously shown LiCMS could gain allosteric sensitivity to leucine with a single mutation: we

hoped that using similar methods we could engineer isoleucine sensitivity into MtuIPMS. A range of single mutants had been tested[1] with partial success: a gain in norvaline sensitivity in one mutant (A567V, which corresponds to the V468A mutation Zhang *et al.* found to confer leucine sensitivity on LiCMS). We hoped by incorporating more mutations we might also achieve an isoleucine-sensitive α-IPMS. One of our enzymes with two mutations (A567V + I627A) did show strong inhibition in the presence of isoleucine, and a corresponding decrease in leucine sensitivity. Two more enzymes (A567V + V551L, and A567V + V551L + I627A) had become allosterically promiscuous, being inhibited by a range of branched-chain amino acids. This was achieved with only two or three mutations to the wild-type enzyme, which may reflect the kinds of change we would see in enzyme evolution, hinting at the close evolutionary relationship between the CMS and α-IPMS enzymes.

Having achieved some insight into the residues mediating molecular recognition at the allosteric binding pocket, we were also interested in exploring the signal-transduction pathway between the allosteric and catalytic sites. This has been investigated by De Carvalho, Frantom, and colleagues[24,29,53,54] by mutagenesis and hydrogen-deuterium exchange experiments.

Koon *et al.*[9] suggested the highly conserved Tyr410 residue in *Mtu*IPMS had a catalytically important function, orientating His379 to act as a base in catalysis. Frantom *et al.*[24] found Tyr410 to be essential for allosteric regulation in *Mtu*IPMS, as substitution with phenylalanine mimicked a completely inhibited state in the enzyme, with a large decrease in catalytic rate that was unaffected by leucine.

We found that the corresponding residue in *Nme*IPMS (Tyr313) was indeed critical for catalysis, as substitution with alanine severely curtailed enzymatic activity. However, when phenylalanine was substituted instead, the enzyme still

showed sensitivity towards leucine: moderate inhibition at low concentrations, and strong activation at higher concentrations. We propose that the mutation compromises enzyme stability leading to reduced catalysis, but is somewhat stabilised (and catalysis rescued) by higher concentrations of leucine. Assessing the thermal stability of this mutant with varying concentrations of leucine ought to provide a simple test of this hypothesis.

The approach adopted by Frantom *et al.* provides excellent experimental insight into these systems, but little predictive power. MD simulations afford many opportunities to generate hypotheses for experimental testing. In this study we tested the properties of several new residues identified by MD simulations as being important for allostery. These residues had not been previously identified as they are not highly conserved, nor do they occupy any obvious structural role.

Two such residues, Arg32 and Asp375, are predicted to form a salt-bridge in one conformation of the asymmetric dimer, potentially affecting stability, catalysis, or inhibition. By substituting each residue for cysteine, we found that the salt-bridge does not appear to be important for catalysis, but does appear to be part of the allosteric communication pathway between regulatory and catalytic domains. We would expect a double-mutant to form a disulfide bridge, restricting movement between domains. The effect this might have on catalysis or allostery may help to shed light on the role of domain movements in α-IPMS enzymes.

Another residue predicted to be involved in domain movements, albeit more subtle, was Lys332. MD simulations suggested it was involved in the movement of subdomain I away from the catalytic barrel upon leucine binding. By mutating this residue to alanine, we found evidence that rather than being responsible

for this movement and inhibition, Lys332 appears instead to stabilise subdomain I near the active site, moderating the effects of leucine binding.

The final residue tested, Arg470, was shown to play a key role in transmitting the allosteric signal out of the regulatory domain. MD simulations predicted altered hydrogen-bonding of this residue upon leucine binding, and we found that mutating it to alanine completely eliminated leucine inhibition.

While we have achieved some fascinating insights into the selectivity and mechanism of allostery in two α-IPMS enzymes, full kinetic characterisation of our mutants should provide more robust evidence for what we have found, allowing stronger conclusions to be drawn. Futher investigation into the Arg32 / Asp375 mutants, and construction and testing of a double-mutant, should provide some interesting clues as to the role of domain movements in α-IPMS.

To conclusively test whether a pendulum-like rearrangement occurs, we would need to explore a Förster Resonance Energy Transfer (FRET) based technique,[55] in particular, single-molecule FRET.[56,57] Introducing two different fluorescent labels onto the dimer could prove difficult. Mutating surface residues to cysteine would allow us to add the labels but we would have to find a way to selectively protect and deprotect these residues in order to bind separate fluorescent donor and acceptor groups.[58] Cell-free translation systems offer greater control[59] but require elaborate preparations. Perhaps the best approach would be to express the N- and C-terminal regions of *Nme*IPMS separately, label them, then use chemical ligation[60] to reconstitute the full-length enzyme. The labelling steps would have to be conducted with a stoichiometric excess of enzyme, to minimise the number of enzymes with both chains labelled.

Gaining a better understanding of the catalytic and regulatory functions of enzymes have diverse possible applications: improving enzymes for green chemical processes,[61–64] as biosensors,[13,65,66] and designing enzyme inhibitors

for drug development.[12,67] The approaches followed here are simple, and yet show the power we have to alter enzymes to suit our needs. This dynamic and emerging field is increasing using a combination of rational models, sequence-based data (including "smart" methods such as ProSAR[68]), and MD simulations[69–71] to open up exciting new avenues to improved therapeutics, and cleaner, more efficient chemistry.

Chapter 5

Materials and Methods

5.1 General Methods

5.1.1 Water

All buffers and solutions were made using reverse-osmosis water treated with a Millipore Milli-Q system prior to use.

5.1.2 pH

The pH of buffers and solutions was measured using a Mettler Toledo Seven-Compact pH meter, and adjusted by addition of NaOH or HCl.

5.1.3 Concentration Determination of DNA & Proteins

Concentration of DNA was measured spectrophotometrically at 260 nm using a Nanodrop® ND-1000 spectrophotometer. Concentration of protein was established similarly at 280 nm, using extinction coefficients from Section 5.5.1.

5.1.4 Protein Structure Images

Protein structures were viewed, and images prepared using The PyMOL Molecular Graphics System, Version 1.5.0.4 Schrödinger, LLC.[72]

5.1.5 Multiple Sequence Alignments

Multiple sequence alignments were created using Clustal Omega[46] and Jalview.[45]

5.2 Microbiology

5.2.1 Cell Lines

E. coli OneShot® TOP10 (Life Technologies), DH5α™(Life Technologies), XL1-Blue (Stratagene), or Stellar™ Competent Cells (Clontech) were used for plasmid propagation. For protein expression, *E. coli* OneShot® BL21 Star (DE3) (Life Technologies) cells were used for *Nme*IPMS proteins and Rosetta 2 (Novagen®) cells for *Mtu*IPMS proteins.

5.2.2 Glycerol Stocks

Cell lines of interest were preserved as glycerol stocks at -80°C. A 5 mL LB culture containing appropriate antibiotics, innoculated from a single colony from an agar plate, was grown overnight, and 0.8 mL of this was added to 0.4 mL sterile 50% glycerol solution (molecular biology grade) in a sterile microcentrifuge tube. After inverting to mix, the tube was labelled and frozen in liquid nitrogen before storage.

5.2.3 Culture Media

LB was used as the standard growth medium for expression and propagation strains, while SOC media was used for transformations. All media was made up in Milli-Q and sterilised in an autoclave prior to use.

Lysogeny Broth (LB)

Compound	Amount
Tryptone	10 g/L
Yeast extract	5 g/L
NaCl	5 g/L

For LB-agar, 15 g/L agar was added prior to autoclaving. LB-agar was melting in a microwave and cooled until approximately 50°C. In a sterile laminar flow cabinet, antibiotics were added as appropriate, and the medium was poured into five or six sterile plastic petri dishes. Once cooled, the plates were used straight away, or wrapped in plastic and stored at 4°C for up to two weeks before use. Stored plates were warmed at 37°C before use.

Super Optimal Broth with Catabolite Repression (SOC) Media

Compound	Amount
Tryptone	20 g/L
Yeast extract	5 g/L
NaCl	10 mM
KCl	2.5 mM
$MgSO_4$	10 mM
Glucose	20 mM

Autoclaved SOC was aliquoted into sterile 1.7 mL microcentrifuge tubes and stored at 4°C until use.

5.2.4 Antibiotics

All plasmids used in this project contain an ampicillin-resistance gene, so all cultures were grown in the presence of 100 μg/mL ampicillin. Rosetta 2 cells are chloramphenicol resistant, so also had 25 μg/mL chloramphenicol added. Antibiotics were stored in sterile microcentrifuge tubes at -20°C or -80°C as a 1000x stock. Chloramphenicol stock was made up in 100% AnalaR ethanol, while ampicillin was made up in either Milli-Q or 50% AnalaR ethanol.

5.3 Genetic Manipulation

5.3.1 Plasmid Extraction

Plasmids were extracted from overnight LB cultures containing appropriate antibiotics, innoculated from either a single colony from an agar plate, or from a

frozen glycerol stock scraping. Approximately 3 mL of the resulting culture was processed using either a NucleoSpin® Plasmid EasyPure kit (Machery-Nagel) or High Pure Plasmid Isolation Kit (Roche). Plasmids were assessed for purity by 260 nm/280 nm absorption measurement and by agarose gel electrophoresis (see section 5.3.3) before being stored at -20°C.

5.3.2 Chemical Transformation

Chemically Competent Cells

To make chemically competent cells, a 5 mL overnight LB culture (with antibiotics as appropriate) was innoculated with a frozen glycerol stock scraping and incubated at 37°C. Around 2 mL of this culture was then used to innoculate 100 mL LB (and antibiotics), which was then grown until an $OD_{600} \sim 0.4$ was reached. The culture was then chilled on ice for at least 30 minutes. Cells were harvested at 3000 rpm (4°C) for 15 minutes in chilled 50 mL sterile centrifuge tubes, and the supernatant discarded. The cells were then kept on ice, and resuspended in 20 mL of chilled, autoclave-sterilised 50 mM $CaCl_2$ + 15% (v/v) glycerol. The cells were incubated on ice for 30 minutes before being harvested again at 3000g as before. The supernatant was discarded and the pellet resuspended on ice in 4 mL $CaCl_2$/glycerol before being aliquoted out into sterile, chilled 1.5 mL centrifuge tubes and frozen in liquid nitrogen. Cells were stored at -80°C until required.

Transformation of Competent Cells

Aliquots of cells were thawed on ice, then up to 5 μL of plasmid or DpnI-digested SDM product was gently mixed in, and the incubated for 30 minutes on ice. The tube was then heat-shocked for 30s in a 42°C waterbath, chilled on ice briefly,

then incubated at 37°C with 500 μL SOC medium. This was then plated onto LB-agar plates with appropriate antibiotics and grown at 37°C.

5.3.3 Agarose Gel Electrophoresis

Agarose gels (1%) were prepared by adding 0.3 g molecular biology-grade agarose in 30 mL TAE buffer and microwaving until dissolved. After allowing several minutes to cool, 3 uL SYBR® Safe was added and the gel poured. Samples were mixed on a piece of Parafilm® M with 6x DNA loading buffer before being loaded onto the gel, which was then run in TAE buffer at 85-100 V for 30-45 minutes in a Mini-Sub® Cell GT (Bio-Rad). Gels were visualised under UV light in a Molecular Imager® Gel Doc™ XR (Bio-Rad).

TAE Buffer (50x)

Compound	Amount
Tris Base	242g
Disodium EDTA	18.6 g
Glacial acetic acid	57.1 mL

Made up to 1 L in Milli-Q

DNA Loading Buffer (6x)

Compound	Amount
Tris-HCl	60 mM
EDTA	60 mM
Orange G	0.2% (w/v)
Xylene Cyanol FF	0.05% (w/v)
Glycerol	60% (v/v)

5.3.4 Polymerase Chain Reaction

Taq DNA Polymerase (Roche) was sometimes used to check the integrity of mutated plasmids or reliability of primers before sequencing. 20-50 μL reactions were prepared according to manufacturer's instructions, and thermocycling

carried out in either a Veriti® 96-well Thermal Cycler (Applied Biosystems) or an iCycler (Bio-Rad). For *Mtu*IPMS-derived templates, 3% DMSO was typically added to the reaction to help overcome the high G+C content typical of *Mycobacterium* genes.

5.3.5 Site-Directed Mutagenesis (SDM)

SDM was typically carried out in 20-50uL reactions using Phusion® High-Fidelity DNA Polymerase (New England BioLabs Inc.®). Around 100ng template plasmid was used per reaction, 5pmol of each forward and reverse mutagenesis primer, and dNTPs etc. as per manufacturer's instructions. For *Nme*IPMS templates the supplied HF buffer was used, whereas for *Mtu*IPMS the GC buffer and 3% DMSO were helpful in overcoming the high G+C content. Thermocycling was as per manufacturer's instructions, using an annealing temperature of 55°C. The SDM product was incubated for 1-2 hours at 37°C in the presence of Dpn I enzyme (either Roche, Life Technologies, or New England BioLabs Inc.®) to digest the template. The resultant digest was transformed into either TOP10 or Stellar Competent Cells for plasmid propagation. Several colonies were picked for plasmid extraction, and the resultant plasmids and their PCR products were typically checked by agarose gel electrophoresis before being sent for sequencing.

5.3.6 DNA Sequencing

Sequencing was performed by either Canterbury Sequencing or Massey Genome Services. Canterbury use an ABI3100 Genetic Analyzer (Applied Systems Inc.) and a procedure based on Sanger chain-termination protocol. Plasmids were provided submitted at 100ng/uL and primers at 3.2 μM. Massey use an ABI3730 DNA Analyzer and BigDye™ Terminator chemistry. Plasmid and template were

provided pre-mixed in 20 μL volumes containing 350-600ng plasmid and 4pmol primer.

5.3.7 Primers for Sequencing & Mutagenesis

Primers for mutagenesis were designed using PrimerX[73] and synthesised by Life Technologies. Primers were delivered lyophilised, and were reconstituted and stored in TE buffer (10 mM Tris-HCl, 100 mM EDTA). A full list of primers used is found in Table Table 5.1.

Table 5.1: Sequencing and SDM Primers

Sequencing	
*Nme*IPMS-FORWARD	TTAATACGACTCACTATAGGG
*Nme*IPMS-REVERSE	TAGTTATTGCTCAGCGGTGGCAGC
*Mtu*IPMS-FORWARD	AGCGGATAACAATTTCACACA
*Mtu*IPMS-INTERNAL-FORWARD	CTGCTGGTCCGGATGGGCTACA
*Mtu*IPMS-REVERSE	ATCTGTATCAGGCTGAAAATC
Mutagenesis	
*Mtu*IPMS-A567V-forward	GACGACGCTCAGGTGGCCGCGTATGTG
*Mtu*IPMS-A567V-reverse	CACATACGCGGCCACCTGAGCGTCGTC
*Mtu*IPMS-A558I-forward	GACTACTACGAGCACATTATGAGCGCCGGCGAC
*Mtu*IPMS-A558I-reverse	GTCGCCGGCGCTCATAATGTGCTCGTAGTAGTC
*Mtu*IPMS-V551L-forward	GTTTGACGTGGCCCTGCTGGACTACTAC
*Mtu*IPMS-V551L-reverse	GTAGTAGTCCAGCAGGGCCACGTCAAAC
*Mtu*IPMS-I627A-forward	GTATCGCACCGTCAGCGACCACCGCGTCGC
*Mtu*IPMS-I627A-reverse	GCGACGCGGTGGTCGCTGACGGTGCGATAC
A567V-FWD-II	GACGACGCTCAGGTAGCCGCTTATGTGGAGG
A567V-REV-II	CTCCACATAAGCGGCTACCTGAGCGTCGTCG
A567V-FWD-III	CGGCGACGACGCTCAGGTAGCCGCTTATGTGGAGG
A567V-REV-III	CCTCCACATACGCGGCTACCTGAGCGTCGTCGCCG
D375A-FWD	CAAAAAACGCGAAATCTTCGCGGAAGACCTGCACGCACTGG
D375A-REV	CCAGTGCGTGCAGGTCTTCCGCGAAGATTTCGCGTTTTTTG
D375C-FWD	CAAAAAACGCGAAATCTTCTGCGAAGACCTGCACGCACTG
D375C-REV	CAGTGCGTGCAGGTCTTCGCAGAAGATTTCGCGTTTTTTG
R32A-FWD	CAAAGAGGAAAAAATCGCCGTCGCCCGCCAGCTGG
R32A-REV	CCAGCTGGCGGGCGACGGCGATTTTTTCCTCTTTG
R32C-FWD	CCAAAGAGGAAAAAATCTGCGTCGCCCGCCAGCTG
R32C-REV	CAGCTGGCGGGCGACGCAGATTTTTTCCTCTTTGG
K332A-FWD	CCGTTTGAGCTTGGGCGCGTTGTCCGGCCGCAACG
K332A-REV	CGTTGCGGCCGGACAACGCGCCCAAGCTCAAACGG
R470A-FWD	GGGCGAAACCAGCGTCGCGCTGGCGCGCGGCAACC
R470A-REV	GGTTGCCGCGCGCCAGCGCGACGCTGGTTTCGCCC
T341K-FWD	CCGCAACGCCTTCAAAAAAAAGCTGGCGGATTTGG
T341K-REV	CCAAATCCGCCAGCTTTTTTTTTGAAGGCGTTGCGG
Y313F-FWD	GAAACACCGCGAAACTTTTGAGATTATGTCCGCCG
Y313F-REV	CGGCGGACATAATCTCAAAAGTTTCGCGGTGTTTC
Y313A-FWD	GAAACACCGCGAAACTGCGGAGATTATGTCCGCCG
Y313A-REV	CGGCGGACATAATCTCCGCAGTTTCGCGGTGTTTC

5.4 Protein Expression & Purification

5.4.1 Protein Expression

A scraping from a glycerol stock was used to innoculate 5 mL LB + ampicillin preculture, which was grown overnight. This was then used to innoculate (typically) 1L of LB + ampicillin, which was then grown at 37°C until an OD_{600} of 0.4-1.0 was reached. Expression was induced with 0.25 mM IPTG and the culture left to grow at 23°C overnight. Cells were typically harvested by centrifugation at 14,000 g for 30 minutes at 4°C before being frozen and stored at -80°C.

5.4.2 Protein Extraction

Cell pellets were thawed at room temperature and resuspended in chilled, 0.2 μM filtered lysis buffer (50 mM potassium phosphate, 300 mM KCl, pH 8.0). Usually around 25 mL of buffer was used for a pellet from 1L culture. Cells were lysed using an Omni-Ruptor 4000 Ultrasonic Homogenizer (Omni International). Cells were kept on ice during sonication, which was performed at 70% power, 40% duty cycle, for cycles of 5 minutes. Cells were allowed to cool between sonications. Sonication progress was monitored by eye - usually the consistency was noticeably thinner after 3-4 cycles. At this stage, Benzonase® Nuclease (Sigma) was added and the soluble protein fraction was separated by centrifugation (40,000 g, 30 minutes at 4°C). If lysate was to be kept for over an hour, one cOmplete Protease Inhibitor Cocktail Tablet (Roche) was added per 50 mL lysate.

5.4.3 FPLC Chromatography

All chromatography was performed using a Bio-Rad BioLogic DuoFlow, either at 4°C or on the bench using ice-chilled buffers and collection tubes in an ice-bath. Elution of protein was followed by UV absorption (benchtop unit) or 280 nm absorption (4°C unit). Samples were typically injected onto columns using a 10 mL or 50 mL Superloop™ (GE Healthcare).

5.4.4 Immobilised Metal Affinity Chromatography (IMAC) Separation of Tagged Protein

Crude lysate was loaded onto a 5 mL Histrap HP (GE Healthcare) IMAC column equilibrated with lysis buffer. Using the Duoflow units' buffer mixing capability, imidazole concentrations could be varied between 0-500 mM using lysis buffer on one pump and elution buffer (lysis buffer + 500 mM imidazole, pH adjusted to 8.0) on the other. The column was washed with 10 column volumes of 20 mM imidazole to remove any non-specificly bound protein. Bound protein was then eluted using an imidazole concentration gradient increasing to 500 mM imidazole over ten column volumes.

5.4.5 Sample Concentration, Desalting, & Buffer Exchange

For concentrating and desalting or buffer exchanging protein samples, 10kDa centrifugation concentrators (GE Healthcare, Sartorius, or Millipore) were used in centrifuges cooled to 4°C. For larger volumes or faster buffer exchanging / desalting, a HiPrep 26/10 Desalting chromatography was used (GE Healthcare) equilibrated with the appropriate final buffer.

5.4.6 His-Tag Cleavage

To remove the His-tag, protein samples were desalted or buffer exchanged to remove imidazole, then 1 mM DTT and 0.5 mM EDTA were added, along with 0.1-1 mg of TEV protease per 10 mg tagged protein. Samples were typically incubated on the bench for 30-60 minutes then overnight at 4°C. Treated samples were then desalted to remove EDTA and loaded onto a HisTrap. Cleaved protein elutes in the wash step, while uncleaved protein, cleaved tags, and TEV protease remain bound to the column.

TEV Protease Purification

Recombinant TEV protease[74] was purified in a similar method to the other enzymes in this work. BL21 (DE3) Star cells transformed with pRIL (rare-codon production) and pRK793 (expressing TEV protease) were grown and expressed in LB with chloramphenicol and kanamycin as above. Extraction and HisTrap separation was performed as above, except 500 mM NaCl was used in buffers instead of 300 mM KCl. Eluted protein was buffer-exchanged into storage buffer (50 mM potassium phosphate, 200 mM KCl, 10 mM DTT, 2 mM EDTA, pH 8.0), concentrated to $\approx$ 1.5 mg/mL, and glycerol added to 10% (v/v), before being flash frozen in aliquots of 1 mg total protein.

5.4.7 Size-Exclusion Chromatography

As a final purification step, samples were loaded onto a Superdex 200 26/60 column (GE Healthcare) equilibrated with 10 mM Tris (pH 8.0) and eluted at 1.0 mL/min. Protein-containing fractions were checked for purity by SDS-PAGE before being pooled and stored.

5.4.8 Protein Storage

Pure protein samples were concentrated to ≥ 1 mg/mL before being flash frozen in small aliquots using liquid nitrogen. Labelled samples were then stored at -80°C until required.

5.5 Protein Characterisation

5.5.1 Physical Parameters

The parameters displayed in Table Table 5.2 were calculated using the ProtParam[75] tool on the Expasy web server. The *Mtu*IPMS mutants included an N-terminal gly-ala remnant from the His-tag. *Nme*IPMS proved very resistant to His-tag cleavage, so the wild-type was purified in both cleaved and tagged form, but the others were only purified with the tag intact. The tagged enzymes had MHHHHHHGKPIPNPLLGLDSTENLYFQ**GA** (*Mtu*IPMS) or MHHHHHHGKPIPNPLLGLDSTENLYFQ**GIDPFT** (*Nme*IPMS) before their sequence, while the cleaved WT protein had only the residues indicated in bold.

Table 5.2: Calculated Protein Parameters

Enzyme	Molecular Weight (Da)	Extinction Coefficient ($M^{-1}cm^{-1}$)
*Mtu*IPMS wild-type	70,242	72,560
Mtu-1	70,312	”
Mtu-2	70,284	”
Mtu-3	70,326	”
Mtu-4	70,228	”
Mtu-5	70,270	”
Mtu-6	70,242	”
Mtu-7	70,284	”
*Nme*IPMS wild-type	56,027	24,660
*Nme*IPMS wild-type (tagged)	59,180	26,150
Y313A*	59,088	”
Y313F*	59,164	”
D375C*	59,168	”
R32C*	59,127	”
K332A*	59,123	”
R470A*	59,095	”

* His-tag not cleaved.

5.5.2 Polyacrylamide Gel Electrophoresis (SDS-PAGE)

SDS-PAGE was used during purification to check protein purity. Samples were mixed with Bolt® sample loading buffer and DTT then boiled for 5 minutes, then loaded onto Bolt® Bis-Tris Plus precast gels (Life Technologies). Gels were run in Bolt® MES or MOPS SDS Running Buffer in a Mini Gel Tank (Life Technologies) at 165-200V for 30-45 minutes. Gels were then heated in a microwave oven with stain (1% Coomassie Brilliant Blue R-250, 40% methanol, 10% acetic acid) and shaken for 5-10 minutes before being destained (40% methanol, 10% acetic acid) with heating and shaking as before, until the background stain had been sufficiently reduced. Gels were stored in RO water until ready for photographing on a Molecular Imager Gel Doc XR (Bio-Rad).

5.5.3 Mass Spectrometry

Samples were desalted into either Milli-Q water or 5 mM ammonium bicarbonate using centrifugation concentrators (Vivaspin® by Vivaproducts or Nanosep®by PALL®) and analysed using Bruker maXis 3G mass spectrometer by the University of Canterbury Chemistry Department.

5.5.4 Circular Dichroism Spectrophotometry

Samples were analysed using a JASCO J-815 Spectropolarimeter. Samples were diluted in a 1 cm pathlength, 3 mL quartz cuvette with Milli-Q water to approximately 0.05 mg/mL and this was adjusted as necessary to achieve the best data signal. Samples were scanned from 190 - 260 nm with 2 nm pitch, 1 second response, and 1 mM bandwidth, at room temperature. Spectra were smoothed and a blank (pure Milli-Q) was subtracted using the supplied JASCO software. Data were scaled before being overlayed to compensate for slightly varying concentrations used.

5.5.5 Differential Scanning Fluorimetry

Differential Scanning Fluorimetry was carried out in a Bio-Rad iCycler™ Real-Time PCR Detection System. Samples were made up in triplicate 25 μL volumes prepared in a 96-well plate on ice. Before analysis, the plate was sealed and centrifuged briefly, then analysed using a thermal melt program on the iCycler® starting at 20°C and increasing by 0.5°C steps (with a 30s hold time) up to 95°C. Control wells were also analysed with all sample components except protein, which was substituted with Milli-Q. Sample components listed below.

Assay Component	Volume (μL)
50 mM BTP (pH 8.0)	20
250x SYPRO® Orange dye (Sigma)	1
Protein sample (1 mg/mL)	4
Total	**25**

*Buffer containing 0.0 mM, 0.5 mM, or 5.0 mM of either leucine, isoleucine, valine, or norvaline.

5.6 Enzyme Kinetics

5.6.1 Equipment

Kinetic assays were performed using a Varian Cary 100 UV-Vis spectrophotometer fitted with temperature controller, and the Cary WinUV Kinetics Application (Version 3, Varian). Initial rates were measured using this application by least-squares fit. Either 2 mM pathlength or black-walled 1 cm pathlength quartz cuvettes were used for all assays. All assays were carried out at 25°C.

5.6.2 Coupled Assay

Adapted from the method from de Carvalho *et al*,[49] the formation of CoA was followed by its reaction with DTP, producing a species which absorbs at 324 nm (Beer-Lambert Law, $A = \varepsilon cl$, where $\varepsilon = 1.98 \times 10^4$ M^{-1} cm^{-1}). A typical assay would include 0-250 μM AcCoA, 0-250 μM α-KIV, and 100 μM DTP made up in 1 mL of assay buffer (50 mM HEPES, 20 μM KCl, 20 μM $MgCl_2$, pH 7.5) brought to temperature and initiated by addition of enzyme. Enzyme concentrations were usually $\approx$ 100 nm, resulting in reaction rates $\leq$ 0.1 Abs/min. Reaction rates were converted from Absorbance units/min to μM/min via the Beer-Lambert Law. Reactions were carried out in duplicate, and repeated if variance greater than 10% between replicates was observed.

Assay Buffer	Concentration	Assay Component	Concentration
HEPES	50 mM	AcCoA	0-250 μM
KCl	20 mM	α-KIV	0-250 μM
$MgCl_2$	20 mM	DTP	100 μM
pH	7.5	Assay buffer	Made up to 1 mL

Substrate Concentration Determination

To find the concentration of AcCoA and α-KIV, a limiting amount of the substrate of interest was added to a reaction cuvette, with the other substrate in excess. The total change in absorbance seen over the course of the reaction ($\Delta A_{reaction}$) was measured in triplicate. The change in absorbance after adding enzyme to the assay buffer with no substrate present was measured as a control (ΔA_{enzyme}). The corrected change in absorbance ($\Delta A_{reaction} - \Delta A_{enzyme}$) was

used with the extinction coefficient ($\varepsilon = 1.98 \times 10^4$ M^{-1} cm^{-1}) and the Beer-Lambert Law to determine the concentration of the limiting substrate.

Inhibition Studies

For inhibition assays, the reaction was initiated by addition of AcCoA rather than enzyme. In this case the reaction mixture (including enzyme, α-KIV, and the amino acid being tested for inhibition) was incubated at 25°C before initiation, to control for the slow-onset inhibition seen in this enzyme. Amino acid concentration was varied between 0, 0.5, 1.0, and 5.0 mM. The reduction in activity was calculated as $(Rate_0 - Rate_i)/(Rate_0) \times 100\%$, where $Rate_0$ is the rate with 0 mM amino acid added and $Rate_i$ is the rate with a given concentration of amino acid present. AcCoA and α-KIV were kept at saturating concentrations to ensure that small volume measurement errors did not significantly affect the rates. Concentrations of 250 μM α-KIV and 425 μM AcCoA were empirically determined to provide a satisfactory balance between economy and reproducibility. Due to the high concentration of AcCoA required, these experiments were conducted using 2 mM pathlength cuvettes.

Bibliography

[1] Clarke, T. Studies on the inhibitor selectivity and inhibitory signal transfer of α-Isopropylmalate synthase. Thesis, University of Canterbury, 2013.

[2] WHO, *Antimicrobial Resistance Global Report on Surveillance: summary*; 2014; p 8.

[3] Kumarasamy, K. K. et al. *The Lancet Infectious Diseases* **2010**, *10*, 597–602.

[4] Freire-Moran, L.; Aronsson, B.; Manz, C.; Gyssens, I. C.; So, A. D.; Monnet, D. L.; Cars, O. *Drug Resistance Updates* **2011**, *14*, 118–124.

[5] Koul, A.; Arnoult, E.; Lounis, N.; Guillemont, J.; Andries, K. *Nature* **2011**, *469*, 483–490.

[6] WHO, *Global tuberculosis report 2014 (WHO/HTM/TB/2014.08)*; 2014.

[7] Zumla, A. *The Lancet* **2011**, *377*, 10–11.

[8] Grandoni, J. a.; Marta, P. T.; Schloss, J. V. *Journal of Antimicrobial Chemotherapy* **1998**, *42*, 475–482.

[9] Koon, N.; Squire, C. J.; Baker, E. N. *Proceedings of the National Academy of Sciences of the United States of America* **2004**, *101*, 8295–8300.

[10] Bange, F. C.; Brown, A. M.; Jacobs, W. R. *Infection and Immunity* **1996**, *64*, 1794–1799.

[11] Hondalus, M. K.; Bardarov, S.; Russell, R.; Chan, J.; Jacobs, W. R.; Bloom, B. R. *Infection and Immunity* **2000**, *68*, 2888–2898.

[12] Nussinov, R.; Tsai, C.-J. *Cell* **2013**, *153*, 293–305.

[13] Villaverde, A. *FEBS Letters* **2003**, *554*, 169–172.

[14] Gallagher, D. T.; Gilliland, G. L.; Xiao, G.; Zondlo, J.; Fisher, K. E.; Chinchilla, D.; Eisenstein, E. *Structure (London, England : 1993)* **1998**, *6*, 465–75.

[15] Zhang, P.; Ma, J.; Zhang, Z.; Zha, M.; Xu, H.; Zhao, G.; Ding, J. *The Biochemical Journal* **2009**, *421*, 133–143.

[16] McCourt, J. a.; Duggleby, R. G. *Amino Acids* **2006**, *31*, 173–210.

[17] Qian, J.; Khandogin, J.; West, A. H.; Cook, P. F. *Biochemistry* **2008**, *47*, 6851–8.

[18] Mulholland, A. J.; Richards, W. *Journal of Molecular Structure: THEOCHEM* **1998**, *429*, 13–21.

[19] de Kraker, J.-W.; Gershenzon, J. *The Plant cell* **2011**, *23*, 38–53.

[20] Quartararo, C. E.; Blanchard, J. S. *Biochemistry* **2011**, *50*, 6879–87.

[21] Chanchaem, W.; Palittapongarnpim, P. *Tuberculosis (Edinburgh, Scotland)* **2002**, *82*, 1–6.

[22] Chanchaem, W.; Palittapongarnpim, P. *FEMS Microbiology Letters* **2008**, *286*, 166–170.

[23] Zhang, Z.; Wu, J.; Lin, W.; Wang, J.; Yan, H.; Zhao, W.; Ma, J.; Ding, J.; Zhang, P.; Zhao, G.-P. *The Journal of Biological Chemistry* **2014**,

[24] Frantom, P. a.; Zhang, H. M.; Emmett, M. R.; Marshall, A. G.; Blanchard, J. S. *Biochemistry* **2009**, *48*, 7457–7464.

[25] Huisman, F. H. a.; Koon, N.; Bulloch, E. M. M.; Baker, H. M.; Baker, E. N.; Squire, C. J.; Parker, E. J. *Biochemistry* **2012**, *51*, 2289–2297.

[26] Webster, R. E.; Gross, S. R. *Biochemistry* **1965**, *4*, 2309–2318.

[27] de Carvalho, L. P. S.; Blanchard, J. S. *Archives of Biochemistry and Biophysics* **2006**, *451*, 141–148.

[28] Ulm, E. H.; Böhme, R.; Kohlhaw, G. *Journal of Bacteriology* **1972**, *110*, 1118–1126.

[29] De Carvalho, L. P. S.; Frantom, P. a.; Argyrou, A.; Blanchard, J. S. *Biochemistry* **2009**, *48*, 1996–2004.

[30] de Kraker, J.-W.; Luck, K.; Textor, S.; Tokuhisa, J. G.; Gershenzon, J. *Plant physiology* **2007**, *143*, 970–86.

[31] De Carvalho, L. P. S.; Blanchard, J. S. *Biochemistry* **2006**, *45*, 8988–8999.

[32] Huisman, F. H. a.; Hunter, M. F. C.; Devenish, S. R. a.; Gerrard, J. a.; Parker, E. J. *Biochemical and Biophysical Research Communications* **2010**, *393*, 168–173.

[33] Kohlhaw, G.; Leary, T. R.; Umbarger, H. E. *Journal of Biological Chemistry* **1969**, *244*, 2218–2225.

[34] Monod, J.; Wyman, J.; Changeux, J. P. *Journal of molecular biology* **1965**, *12*, 88–118.

[35] Koshland, D. E.; Némethy, G.; Filmer, D. *Biochemistry* **1966**, *5*, 365–385.

[36] Cui, Q.; Karplus, M. *Protein science : a publication of the Protein Society* **2008**, *17*, 1295–307.

[37] Goodey, N. M.; Benkovic, S. J. *Nature chemical biology* **2008**, *4*, 474–82.

[38] Xu, H.; Zhang, Y.; Guo, X.; Ren, S.; Staempfli, A. a.; Chiao, J.; Jiang, W.; Zhao, G. *Journal of Bacteriology* **2004**, *186*, 5400–5409.

[39] Atsumi, S.; Liao, J. C. *Applied and environmental microbiology* **2008**, *74*, 7802–8.

[40] Risso, C.; Van Dien, S. J.; Orloff, A.; Lovley, D. R.; Coppi, M. V. *Journal of bacteriology* **2008**, *190*, 2266–74.

[41] Tang, Y. J.; Yi, S.; Zhuang, W.-Q.; Zinder, S. H.; Keasling, J. D.; Alvarez-Cohen, L. *Journal of bacteriology* **2009**, *191*, 5224–31.

[42] Berg, I. A.; Ivanovsky, R. N. *Microbiology* **2009**, *78*, 16–24.

[43] Wu, B.; Zhang, B.; Feng, X.; Rubens, J. R.; Huang, R.; Hicks, L. M.; Pakrasi, H. B.; Tang, Y. J. *Microbiology (Reading, England)* **2010**, *156*, 596–602.

[44] The UniProt Consortium, *Nucleic Acids Research* **2014**, *43*, D204–212.

[45] Waterhouse, A. M.; Procter, J. B.; Martin, D. M. A.; Clamp, M.; Barton, G. J. *Bioinformatics* **2009**, *25*, 1189–1191.

[46] Sievers, F.; Wilm, A.; Dineen, D.; Gibson, T. J.; Karplus, K.; Li, W.; Lopez, R.; McWilliam, H.; Remmert, M.; Söding, J.; Thompson, J. D.; Higgins, D. G. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. 2011.

[47] Niesen, F. H.; Berglund, H.; Vedadi, M. *Nature protocols* **2007**, *2*, 2212–2221.

[48] Tsumoto, K.; Umetsu, M.; Kumagai, I.; Ejima, D.; Philo, J. S.; Arakawa, T. *Biotechnology progress* **2004**, *20*, 1301–1308.

[49] De Carvalho, L. P. S.; Argyrou, A.; Blanchard, J. S. *Journal of the American Chemical Society* **2005**, *127*, 10004–10005.

[50] Nagradova, N. K. *FEBS Letters* **2001**, *487*, 327–332.

[51] Huisman, F. H. a.; Squire, C. J.; Parker, E. J. *Biochemical and Biophysical Research Communications* **2013**, *433*, 249–254.

[52] Huisman, F. H. a. Studies into the allosteric regulation of alpha-isopropylmalate synthase. Thesis, University of Canterbury, 2012.

[53] Frantom, P. a. *Archives of Biochemistry and Biophysics* **2012**, *519*, 202–209.

[54] Casey, A. K.; Schwalm, E. L.; Hays, B. N.; Frantom, P. a. *Biochemistry* **2013**, *52*, 6737–6739.

[55] Weiss, S. *Nature structural biology* **2000**, *7*, 724–729.

[56] Roy, R.; Hohng, S.; Ha, T. *Nature methods* **2008**, *5*, 507–16.

[57] Schuler, B. *Journal of nanobiotechnology* **2013**, *11 Suppl 1*, S2.

[58] Qu, Q.; Sharom, F. J. *Biochemistry* **2001**, *40*, 1413–1422.

[59] Kajihara, D.; Abe, R.; Iijima, I.; Komiyama, C.; Sisido, M.; Hohsaka, T. *Nature methods* **2006**, *3*, 923–929.

[60] Muir, T. W.; Sondhi, D.; Cole, P. a. *Proceedings of the National Academy of Sciences of the United States of America* **1998**, *95*, 6705–6710.

[61] Häring, D.; Schreier, P. *Naturwissenschaften* **1999**, *86*, 307–312.

[62] Sharma, M.; Sharma, N. N.; Bhalla, T. C. *Enzyme and Microbial Technology* **2005**, *37*, 279–294.

[63] Fastrez, J. *ChemBioChem* **2009**, *10*, 2824–2835.

[64] Bornscheuer, U. T.; Huisman, G. W.; Kazlauskas, R. J.; Lutz, S.; Moore, J. C.; Robins, K. *Nature* **2012**, *485*, 185–194.

[65] Zheng, H.; Bi, J.; Krendel, M.; Loh, S. N. *Biochemistry* **2014**,

[66] Stein, V.; Alexandrov, K. *Trends in Biotechnology* **2014**, *33*, 101–110.

[67] Nussinov, R.; Tsai, C. J.; Csermely, P. *Trends in Pharmacological Sciences* **2011**, *32*, 686–693.

[68] Fox, R. J.; Davis, S. C.; Mundorff, E. C.; Newman, L. M.; Gavrilovic, V.; Ma, S. K.; Chung, L. M.; Ching, C.; Tam, S.; Muley, S.; Grate, J.; Gruber, J.; Whitman, J. C.; Sheldon, R. a.; Huisman, G. W. *Nature biotechnology* **2007**, *25*, 338–344.

[69] Durrant, J. D.; McCammon, J. A. *BMC Biology* **2011**, *9*, 71.

[70] Csermely, P.; Korcsmáros, T.; Kiss, H. J. M.; London, G.; Nussinov, R. *Pharmacology and Therapeutics* **2013**, *138*, 333–408.

[71] Milardi, D.; Pappalardo, M. *European Journal of Medicinal Chemistry* **2015**, *91*, 1–3.

[72] Schrodinger LLC, The PyMOL Molecular Graphics System, Version 1.3r1. 2010.

[73] Lapid, C.; Gao, Y. PrimerX. `http://www.bioinformatics.org/primerx`.

[74] Kapust, R. B.; Tözsér, J.; Fox, J. D.; Anderson, D. E.; Cherry, S.; Copeland, T. D.; Waugh, D. S. *Protein engineering* **2001**, *14*, 993–1000.

[75] Gasteiger, E.; Gattiker, A.; Hoogland, C.; Ivanyi, I.; Appel, R. D.; Bairoch, A. *Nucleic Acids Research* **2003**, *31*, 3784–3788.